

# Voice Choice in Äiwoo

## *A Discourse-Functional Study of the Symmetrical Voice System of Äiwoo*

Sindre Sperre Holmen



LING4190  
Master Thesis in Linguistics  
Department of Linguistics and Scandinavian Studies

UNIVERSITY OF OSLO

Autumn 2020



# **Voice Choice in Äiwoo**

A Discourse-Functional Study of the Symmetrical Voice System of Äiwoo

© Sindre Sperre Holmen

2020

Voice Choice in Äiwoo

A Discourse-Functional Study of the Voice System of Äiwoo

Sindre Sperre Holmen

<http://www.duo.uio.no/>

Trykk: Reprosentralen, Universitetet i Oslo

IV

# Abstract

This thesis investigates the motivational factors behind the choice of voice in the Oceanic language Äiwoo. This is a language with a so-called symmetrical voice system, which is characterized by the alternation of two or more clause types. Unlike asymmetrical voice systems, like the active-passive system found in many European languages, the different clause types in a symmetrical voice system are equally marked for voice morphosyntactically and are typically considered equally transitive. While this type of system is found in a number of western Austronesian languages, it is generally considered lost in Proto Oceanic, and most modern Oceanic languages have transitivity-based systems rather than symmetrical voice. Äiwoo, being an Oceanic language, is thus unique among this group of languages in having a symmetrical voice system.

The clause system of Äiwoo consists of three voices: actor voice, undergoer voice and circumstantial voice. The focus of this thesis is the motivation behind the choice of the two main voices, namely actor voice and undergoer voice. In order to investigate this, I have employed three different explanatory models. These are *topic continuity* developed by Givón (1983, 1994), *discourse transitivity* developed by Hopper & Thompson (1980) and the three different *activation states* as discussed by Chafe (1987). I have coded 160 Äiwoo clauses according to the factors described in these models. Through significance testing, I found that aspect and the activation state of undergoers are significant predictors for voice, so that perfective aspect predicts the choice of undergoer voice and new undergoers predict the choice of actor voice.

Based on these results, I argue that undergoer voice corresponds to a higher degree of semantic transitivity than actor voice, a tendency found in other Austronesian languages with symmetrical voice systems. The implications of this give rise to a discussion on the differences between a symmetrical voice system and a transitivity-based system, where I argue that Äiwoo is in the process of developing a transitivity-based system where undergoer voice is reanalysed as the standard transitive clause type.

# Acknowledgements

Det er en lang rekke personer som fortjener en stor takk for at denne masteroppgaven har blitt til. Først og fremst vil jeg takke min hovedveileder, Åshild Næss. Takk for grundige tilbakemeldinger, kloke innspill og ikke minst for å ha vist meg hvor spennende austronesiske språk kan være. Jeg vil også rette en stor takk til min biveileder, Dag Haug, for konstruktiv og grundig hjelp med statistikken.

Det er også rekke venner og medstudenter som fortjener en stor takk. Jeg vil særlig takke Jonas, Marie, Karen og Ingrid for å ha gjort tiden på Henrik Wergelands hus uforglemmelig. Takk også til Joakim som har kommet innom med flytende glede på boks, til Maria for utallige pep-talks, til Brage som har hjulpet med korrekturlesing og til Aafke som har gjort det samme og i tillegg ikke bare har hatt tålmodigheten, men også interessen av å diskutere symmetrisk diatese med meg.

Sist, men ikke minst, takk til Ellisiv og John Harald. Dere har gjort det siste semesteret ikke bare utholdelig, men også utrolig fornøyet! Takk for alle turene til Bunnpris og Dana, de sene kveldene på PC-rommet, for utveksling av memes, og ikke minst for all støtten dere har gitt meg den siste måneden før innlevering!

# Contents

|       |  |    |
|-------|--|----|
| 1     | Introduction .....   | 1  |
| 1.1   | Research question .....  | 1  |
| 1.2   | The Austronesian language family .....                                   | 2  |
| 1.2.1 | Proto Austronesian and Proto Malayo-Polynesian.....                      | 4  |
| 1.2.2 | Oceanic.....   | 6  |
| 1.3   | Symmetrical voice .....  | 8  |
| 1.3.1 | Philippine-type voice system .....                                       | 8  |
| 1.3.2 | Indonesian-type voice system .....                                       | 11 |
| 1.4   | Symmetrical voice in Äiwoo .....   | 14 |
| 1.4.1 | The Äiwoo language .....   | 14 |
| 1.4.2 | Actor voice and undergoer voice .....                                    | 15 |
| 1.4.3 | Circumstantial voice.....  | 18 |
| 1.4.4 | Aspect and mood marking in Äiwoo.....                                    | 20 |
| 1.5   | Symmetrical voice or transitivity alternation?.....                      | 21 |
| 1.6   | Scientific motivations .....   | 28 |
| 1.7   | Outline of thesis.....   | 28 |
| 2     | Previous studies on voice choice in western Austronesian languages ..... | 30 |
| 2.1   | Frameworks .....   | 30 |
| 2.1.1 | Topic continuity .....   | 30 |
| 2.1.2 | Discourse transitivity.....  | 32 |
| 2.1.3 | Activation state.....  | 34 |
| 2.2   | Previous studies on voice choice .....                                   | 35 |
| 2.2.1 | Walters 1994 – Cebuano .....   | 35 |
| 2.2.2 | Wouk 1996 – Spoken Jakartan Indonesian .....                             | 37 |
| 2.2.3 | Wouk 1999 – Sasak.....   | 39 |
| 2.2.4 | Quick 2005 – Pendau .....  | 41 |
| 2.2.5 | McDonnell 2016 – Besemah .....   | 42 |
| 2.2.6 | Summary of results.....  | 44 |
| 2.3   | Previous hypotheses on voice choice in Äiwoo .....                       | 46 |
| 3     | Data and methodology .....   | 47 |
| 3.1   | Data.....  | 47 |

|       |  |    |
|-------|--|----|
| 3.2   | Methodological approach .....              | 49 |
| 3.2.1 | Topic continuity .....                     | 49 |
| 3.2.2 | Discourse transitivity.....                | 51 |
| 3.2.3 | Activation state.....                      | 58 |
| 3.2.4 | Note on terminology.....                   | 60 |
| 4     | Results and analysis .....                 | 62 |
| 4.1   | Results .....                              | 62 |
| 4.1.1 | Topic continuity .....                     | 62 |
| 4.1.2 | Discourse transitivity.....                | 65 |
| 4.1.3 | Activation states .....                    | 74 |
| 4.2   | Significance testing.....                  | 78 |
| 4.3   | Discussion on statistical analysis.....    | 79 |
| 5     | Summary and discussion .....               | 82 |
| 5.1   | Summary.....                               | 82 |
| 5.1.1 | Topic continuity .....                     | 82 |
| 5.1.2 | Discourse transitivity.....                | 83 |
| 5.1.3 | Activation state.....                      | 85 |
| 5.1.4 | Topic continuity vs. activation state..... | 87 |
| 5.2   | Voice choice in Äiwoo .....                | 88 |
| 5.3   | Symmetrical voice vs. transitivity .....   | 89 |
| 5.4   | Future research .....                      | 97 |



# List of abbreviations

|       |                           |      |  |
|-------|---------------------------|------|--|
| 1/2/3 | first/second/third person | MIN  | minimal                                |
| 12    | first and second person   | NEG  | negation                               |
| A     | actor                     | NMLZ | nominalization                         |
| ACC   | accusative                | NPIV | non-pivot                              |
| APLL  | applicative               | O    | undergoer                              |
| APPR  | apprehensive              | OBJ  | object                                 |
| AV    | actor voice               | OBL  | oblique                                |
| AUG   | augmented                 | PART | partitive                              |
| COLL  | collective                | PFV  | perfective                             |
| COM   | comitative                | PL   | plural                                 |
| CONJ  | conjunction               | POSS | possessive                             |
| CORE  | core argument             | PREP | preposition                            |
| CV    | circumstantial voice      | PROX | proximal                               |
| DEF   | definite                  | PV   | patient voice                          |
| DEIC  | deictic                   | RL   | realis                                 |
| DEM   | demonstrative             | S    | single argument of intransitive clause |
| DIR   | directional               | SG   | singular                               |
| DIST  | distal                    | SUBJ | subject                                |
| EXCL  | exclusive                 | THC  | thematic consonant                     |
| FUT   | future tense              | TR   | transitive                             |
| GA    | generic actor             | UA   | unit augmented                         |
| IPFV  | imperfective              | UV   | undergoer voice                        |
| IRR   | irrealis                  | VC   | voice                                  |
| LOC   | locative                  |      |  |
| LV    | locative voice            |      |  |



# 1 Introduction

## 1.1 Research question

The focus of this thesis is the voice system of the Oceanic language Äiwoo. Äiwoo has a system of clause alternations known as a symmetrical voice system. Unlike an asymmetrical voice system, like the active-passive system, a symmetrical voice system consists of two or more clause types that are equally marked for voice morphosyntactically and are typically considered to be equally transitive.<sup>1</sup> The symmetrical voice system of Äiwoo consists of three voices: actor voice, undergoer voice, and circumstantial voice. Exactly what motivates the selection of one voice construction over another in languages with this type of system is widely debated. In this thesis, I investigate which factors motivate the selection of the two main voices in Äiwoo: actor voice and undergoer voice. An example of each of these clause types is shown in (1) and (2).

(1) Actor voice<sup>2</sup>

|                           |                  |                |
|---------------------------|------------------|----------------|
| Pe-sime-engâ              | li-epave=to      | sii=kâ.        |
| COLL-person-DEM:DIST      | 3AUG-cook.AV=now | fish=DEIC:DIST |
| ‘The people cooked fish.’ |                  |                |

(Næss 2015: 274)

(2) Undergoer voice

|                         |           |                                 |
|-------------------------|-----------|---------------------------------|
| Sii                     | lâ        | ki-epavi-i=to=wâ.               |
| fish                    | DEIC.DIST | IPFV-cook.UV-3AUG=now=DEIC:DIST |
| ‘They cooked the fish.’ |           |                                 |

(Næss 2015: 274)

In actor voice, the actor<sup>3</sup> argument, which is *pe-sime-engâ* ‘the people’, is preverbal and the verb takes the form *epave* ‘cook’, signalling actor voice. In undergoer voice, it is the undergoer argument, *sii* ‘fish’, that is preverbal, while the verb takes the form *epavi*, signalling undergoer voice. While actor voice and undergoer voice are equally

---

<sup>1</sup> However, this is debated for some languages (Liao 2004, Ross 2002).

<sup>2</sup> In all the examples in this thesis, I will explicitly state the names of all languages, unless the examples come from Äiwoo. Hence, if the name of the language is not specified, the language is Äiwoo.

<sup>3</sup> In this thesis, I will use the term ‘actor’ to refer to the more agentive argument of a two-argument clause and the term ‘undergoer’ to refer to the more patientive argument of a two-argument clause (Van Valin & LaPolla 1997: 141).

morphosyntactically marked for voice and can be considered to be equally basic, undergoer voice is vastly more frequent than actor voice, with a factor of roughly 7 to 1. I will explain how I found these numbers in more detail in section 3.1.

In this thesis, I will attempt to answer the following question: What factors motivate speakers of Äiwoo to choose one of these voice constructions over the other?

This chapter has the following structure. In section 1.2, I describe the Austronesian language family, of which Äiwoo belongs. Here, I give a geographic overview of where the languages of this family are spoken. I also discuss the basic morphosyntactic structure of the reconstructed languages Proto Austronesian and Proto Malayo-Polynesian (1.2.1) and how the symmetrical voice system found in these languages is thought to have been lost in Proto Oceanic as well as in most languages within the Oceanic group. In section 1.3, I give a more detailed description of symmetrical voice and distinguish between two types of symmetrical voice systems, the Philippine-type system (1.3.1) and the Indonesian-type system (1.3.2). I also discuss how these differ from asymmetrical systems like the active-passive system that is found in most European languages. The symmetrical voice system of Äiwoo is presented in section 1.4. Here, I begin with a brief description of the Äiwoo language and its geographical location (1.4.1), before I discuss the differences between the two main voices in Äiwoo: actor voice and undergoer voice (1.4.2). I also give a brief description of Äiwoo's third voice: circumstantial voice (1.4.3) as well as the aspectual and modal systems of this language (1.4.4). In section 1.5, I discuss whether the clause system of Äiwoo should be analysed as a symmetrical voice system or as a system of transitivity alternations like that found in other Oceanic languages. The term *transitivity discord* is of relevance to this discussion. The scientific motivations behind this thesis are discussed in section 1.6. Finally, I summarize the chapter and present the structure of the rest of the thesis in section 1.7.

## 1.2 The Austronesian language family

The Austronesian language family stretches from Taiwan in the north, to New Zealand in the south, and from Madagascar in the west, to the Easter Island in the east as seen from figure 1, and consists of approximately 1200 languages (Musgrave 2007: 1). Genealogically, the languages are divided into two main groups: the Formosan languages<sup>4</sup>, spoken on Taiwan,

---

<sup>4</sup> The Formosan languages do not actually constitute one single first order subgroup, but several. While the exact number of Formosan groups is debated, Blust (1999) counts nine different first order subgroups.

and the Malayo-Polynesian languages, which constitute the rest of the Austronesian languages. Hence, the majority of languages within the Austronesian family are Malayo-Polynesian. Within the Malayo-Polynesian group, two first-order groups are found: Western Malayo-Polynesian and Central-Eastern Malayo-Polynesian. The Oceanic branch, of which Äiwoo belongs, is part of the Eastern Malayo-Polynesian group within the Central-Eastern Malayo-Polynesian first-order group (Musgrave 2007: 1-2). This overview is summarized in (3). The geographical location and genealogical classification of Äiwoo is discussed in more detail in section 1.4.1.

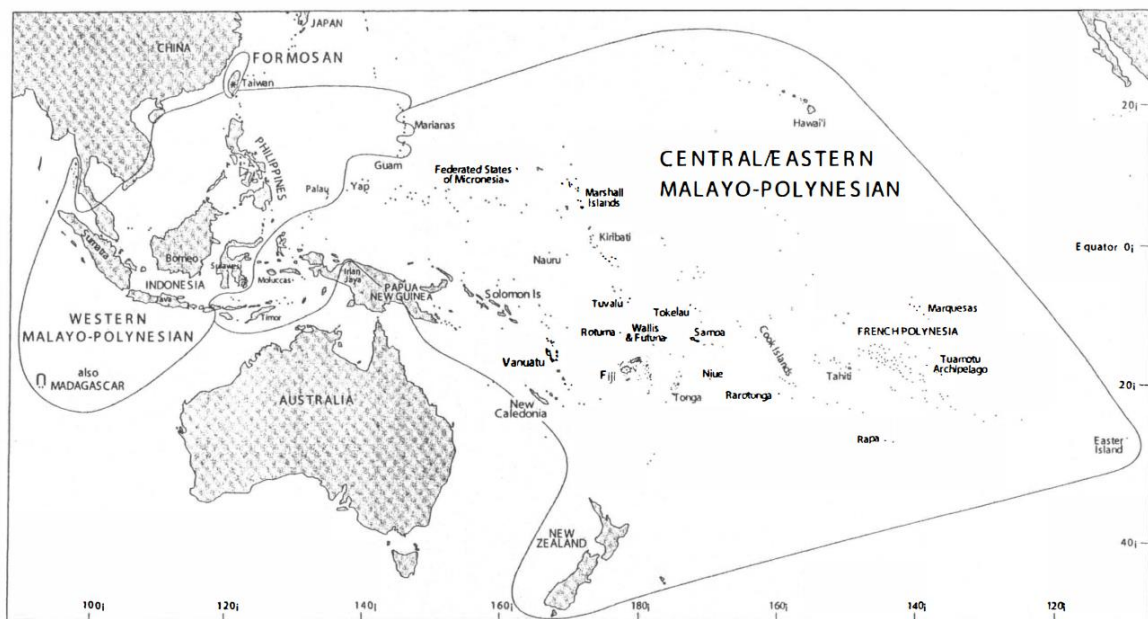


Figure 1.1: The Austronesian language family (Ross 2002: 18)

- (3) Austronesian
- Formosan (several subgroups)
  - Malayo-Polynesian
    - Western Malayo-Polynesian
    - Central-Eastern Malayo-Polynesian
      - Eastern Malayo-Polynesian
      - Oceanic

## 1.2.1 Proto Austronesian and Proto Malayo-Polynesian

Reconstructions of the clause system of Proto Austronesian (henceforth PAn) show that this language had a symmetrical voice system like that found in many western Austronesian languages today. The PAn voice system consisted of an alternation between four different voices: *actor voice*, where the actor argument was the subject<sup>5</sup> of the clause, *patient voice*, where the patient argument was the subject, *location voice*, where the locative argument was the subject, and *circumstantial voice*, where a circumstantial argument like instrument, theme or beneficiary was the subject. Each of these voices were signalled by a specific affix, which is shown in (4).

|     |                              |         |          |                |
|-----|------------------------------|---------|----------|----------------|
| (4) | Actor                        | Patient | Location | Circumstantial |
|     | <um>                         | -ən     | -an      | Si-            |
|     | (Ross 2002: 33) <sup>6</sup> |         |          |                |

These affixes attached to the root of the verb, signalling which of these four voices characterized the clause. Thus, the infix <um> was attached inside the verb root, signalling actor voice, the suffixes -ən and -an were attached to the end of the verb root, signalling patient and location voice respectively, while the prefix Si- was attached to the start of the verb root, signalling circumstantial voice. This system is illustrated in (5) by the examples from Paiwan, a modern Formosan language that has retained this voice system.

|     |        |  |   |         |
|-----|--------|--|---|---------|
|     | Paiwan |  |   |         |
| (5) | a.     | təkəl-ən   | a | vaua    |
|     |        | drink-PV   |   | wine    |
|     |        | ‘the wine will be drunk’ (‘s/he/they will drink the wine’)                     |   |         |
|     | b.     | təkəl-an   | a | kakəsan |
|     |        | drink-LV   |   | kitchen |
|     |        | ‘the kitchen will be drunk in’ (‘s/he/they will drink it/them in the kitchen’) |   |         |

<sup>5</sup> For the sake of terminological simplicity, I use the term subject for the time being. However, this is a somewhat imprecise term, as will be discussed in section 1.3.1 below.

<sup>6</sup> This paradigm is very simplified, showing only the affixes for neutral tense clauses in the indicative mood. The complete paradigm consists of three tenses in indicative mood with the addition of two tenses in non-indicative mood (Ross 2002: 33). However, only the voice distinctions, not the tense or mood distinctions are of relevance here.

- c.     *si-təkəl*         *a*         *kapu*  
           CV-drink                     cup  
           ‘the cup will be drunk with’ (‘s/he/they will drink it/them from a cup’)
- d.     *t<ə>kəl*         *a*         *qala*  
           <AV>drink                     stranger  
           ‘the stranger will drink (something)’
- (Ross 2002: 20-21)

As seen in these examples, *-ən* attaches to the verb root in (5a), signalling patient voice, *-an* attaches to the root in (5b), signalling location voice, *si-* attaches to the root in (5c), signalling circumstantial voice, and finally, *<ə>* is infixes inside the verb root in (5d), signalling actor voice. Notice that these clauses only contain one argument which is preceded by the phrase marker *a*. However, each of these clause types can occur with two arguments. In transitive clauses, i.e. two-argument clauses, the phrase marker *a* precedes the subject and the verb is affixed by the voice affix which corresponds to the semantic relation of the subject. The other argument is preceded by another core argument phrase marker, which is either *nua* or *tua* as in (6).

- (6)   Paiwan
- |                      |          |             |            |             |
|----------------------|----------|-------------|------------|-------------|
| <i>t&lt;ə&gt;kəl</i> | <i>a</i> | <i>qala</i> | <i>tua</i> | <i>vaua</i> |
| <AV>drink            |          | stranger    | NPIV       | wine        |
- ‘the stranger will drink wine’ (Ross 2002: 22)

In (6), the actor, *qala* ‘stranger’ is the subject and is preceded by the phrase marker *a*, while the non-subject core argument *vaua* ‘wine’ is preceded by the non-subject phrase marker *tua*. I discuss this type of voice system in more detail in section 1.3.1, where it becomes clear that not only the actor, but all semantic relations can be subject and consequently be preceded by the “subject” phrase marker *a*, or its corresponding form in different languages.

A similar system has also been reconstructed for Proto Malayo-Polynesian (henceforth PMP). The voice affixes are shown in (7).<sup>7</sup>

---

<sup>7</sup> Again, this paradigm is very simplified.

|     |                 |         |          |                |
|-----|-----------------|---------|----------|----------------|
| (7) | Actor           | Patient | Location | Circumstantial |
|     | <um>            | -ən     | -an      | i-             |
|     | (Ross 2002: 49) |         |          |                |

These affixes would attach to the root of the verb, signalling which voice is chosen in the same way as illustrated by the Paiwan examples above. This type of voice system is attested for a large number of western Austronesian languages. Note that ‘western Austronesian’ is not a genealogical term, but rather a geographical one. It refers to both the Formosan and the Malayo-Polynesian languages that have retained this voice system. An overview of the exact geographical locations of the languages with symmetrical voice systems is provided in section 1.3.

### 1.2.2 Oceanic

Like PAN and PMP, scholars have also reconstructed Proto Oceanic (henceforth POc), which differed significantly from the former two. The development from PMP into POc is a complex process, a detailed discussion of which is beyond the scope of this thesis. The essential thing to note is the differences in the clause systems. POc is reconstructed as having lost the symmetrical voice system found in PAN and PMP. The affixes signalling the different voices in PAN and PMP disappeared, resulting in a simplified clause system characterized by alternations in transitivity, rather than voice (Lynch et al. 2001: 61). This system of transitivity alternations is present in most modern Oceanic languages. Formally, these alternations are signalled by one or two valency-increasing suffixes or enclitics on intransitive verbs, adding an object to the clause and making them transitive. Additionally, an object-marking suffix is also attached to this valency-increasing suffix, signalling the object’s person and number. The person and number of the subject is typically signalled by a preverbal marker of either the S or A argument (Ross 2004: 495). Because the S and A arguments together form a morphosyntactic category distinguished from the O argument in these languages, they have accusative alignment. This is illustrated by the Longgu examples in (8).

|     |    |                          |          |       |
|-----|----|--------------------------|----------|-------|
|     |    | Longgu                   |          |       |
| (8) | a. | Mwaa-i                   | e        | ango  |
|     |    | snake-SG                 | 3SG.SUBJ | crawl |
|     |    | ‘The snake is crawling.’ |          |       |



- b. Mwaa-i e ango-vi-a vanga ngaia  
snake-SG 3SG.SUBJ crawl-TR-3SG.OBJ food 3SG  
‘The snake crawled to/for its food.’
- c. Mwaa-i e ango-ta’ini-ra gale ngaia-gi  
snake-SG 3SG.SUBJ crawl-TR-3PL.OBJ child 3SG-PL  
‘The snake crawled with its babies (on its back).’
- d. Mwaa e ango va’ini-ra gale ngaia-gi  
Snake 3SG.SUBJ crawl COM-3PL.OBJ child 3SG-PL  
‘The snake crawled with its babies (on its back).’ (Hill 1992: 58-60)

(8a) exemplifies an intransitive clause in Longgu. The S argument *mwaa* ‘snake’ is the subject, as indicated by the preverbal subject marker *e*. (8b) and (8c), however, exemplify transitive counterparts of (8a), as signalled by the valency-increasing suffixes *-vi* and *-ta’ini* respectively. These suffixes also indicate which semantic role the object can have.<sup>8</sup> In addition to overt object noun phrases, the object in each clause is furthermore marked on the verb by the suffixes *-a* and *-ra* respectively, indicating the person (third person) and number (singular and plural respectively). The subject in (8b) and (8c) is still *mwaa* ‘snake’ but is now the A argument, not S, because the clauses are transitive rather than intransitive. However, notice that the preverbal subject marker *e* is the same, regardless of whether the subject is S or A. This clearly demonstrates that S and A together form a subject category in Longgu. Finally, (8d) exemplifies an intransitive clause with two participants, the subject *mwaa* ‘snake’ and the oblique argument *gale* ‘babies’. Notice that there is no valency-increasing suffix on the verb in this clause. Instead, the oblique argument is introduced by the preposition *va’ini*. Hence, Longgu clearly distinguishes core from oblique arguments, as is evident when comparing (8c) with (8d). In (8d), *gale* ‘babies’, is an oblique argument as is evident from the presence of the preposition, whereas in (8c) this preposition is absent, indicating that *gale* is a core argument.

In short, because Longgu has clear subject and object categories and makes formal distinctions between core and oblique arguments, its clause system is clearly characterized by

<sup>8</sup> See Hill (1992: 53-61) for an overview of the semantic roles of objects in Longgu.

transitivity alternations rather than voice alternations like that found in PAn and PMP. Now that the typology of the Austronesian language family has been described, I will discuss what is meant by symmetrical voice.

## 1.3 Symmetrical voice

Symmetrical voice refers to a type of system of clause alternations that is found in a large number of western Austronesian languages. This system is also referred to as a focus system, trigger system, or western Austronesian voice system (Chen & McDonnell 2019: 174). This system differs from the asymmetrical voice systems<sup>9</sup>, like the active-passive system found in most European languages. In a symmetrical voice system, two or more clauses alternate according to which semantic argument receives the status of *privileged syntactic argument* (also referred to as subject or pivot) (Chen & McDonnell 2019: 174). The system is *symmetrical* as none of the clauses that alternate is unmarked for voice or considered to be the basic one from which all others are derived. Rather, all the clauses are equally marked for voice morphosyntactically (with the exception of some Indonesian-type languages discussed below). This is contrasted by an asymmetrical voice system, like the active-passive system, where the active is the unmarked, basic clause, while the passive is marked either morphologically or syntactically and is derived from the active (Chen & McDonnell 2019: 175). The symmetrical voice system is divided into two subtypes: the Philippine-type system and the Indonesian-type system, which will be discussed presently.

### 1.3.1 Philippine-type voice system

Philippine-type voice systems typically have a four-way voice distinction consisting of an actor voice, patient voice, locative voice and circumstantial voice (Chen & McDonnell 2019: 176). This type of voice system is found in the Formosan languages spoken in Taiwan (e.g. Paiwan, described in section 1.2.1), as well as the Malayo-Polynesian languages spoken in the Philippines, northern Borneo, and northern Sulawesi. It is also found in Chamorro, spoken on the island Guam, and in Malagasy, spoken on Madagascar, both of which are Malayo-Polynesian languages (Chen & McDonnell 2019: 176). Hence, ‘Philippine-type language’ does not only refer to Philippine languages, but to all languages that have this type of voice

---

<sup>9</sup> See Foley (2008) for a discussion on how a symmetrical voice system differs from both the active-passive system and the ergative-antipassive system.

system (Ross 2002: 20). While the Indonesian-type voice system (to be discussed below) is only attested in Malayo-Polynesian languages, the Philippine-type voice system is attested in both Malayo-Polynesian and Formosan languages. Because of this, the Philippine-type system is reconstructed for both PAN and PMP, as seen in (4) and (7) in section 1.2.1 above. The Paiwan clauses in (5) above illustrated a typical Philippine-type voice system with a four-way voice distinction. Another Philippine-type voice system is illustrated by the Tagalog clauses in (9). Unlike typical Philippine-type languages however, Tagalog has five different voice constructions.

|     |    |   |    |           |      |          |            |                           |
|-----|----|---|----|-----------|------|----------|------------|---------------------------|
|     |    | Tagalog                                   |    |           |      |          |            |                           |
| (9) | a. | <b>b-um-ili</b>                           | ng | isda      | sa   | tindahan | <b>ang</b> | <b>lalake</b>             |
|     |    | VC-buy                                    |    | CORE fish | OBL  | store    |            | man                       |
|     |    | ‘The man bought fish in the store.’       |    |           |      |          |            |                           |
|     | b. | <b>nbi-bilh-in</b>                        | ng | lalake    | sa   | tindahan | <b>ang</b> | <b>isda</b> <sup>10</sup> |
|     |    | IRR-buy-VC                                |    | CORE man  | OBL  | store    |            | fish                      |
|     |    | ‘The man will buy the fish in the store.’ |    |           |      |          |            |                           |
|     | c. | <b>bi-bilh-an</b>                         | ng | lalake    | ng   | isda     | <b>ang</b> | <b>tindahan</b>           |
|     |    | IRR-buy-VC                                |    | CORE man  | CORE | fish     |            | store                     |
|     |    | ‘The man will buy fish in the store.’     |    |           |      |          |            |                           |
|     | d. | <b>ipam-bi-bili</b>                       | ng | lalake    | ng   | isda     | <b>ang</b> | <b>salapi</b>             |
|     |    | VC-IRR-buy                                |    | CORE man  | CORE | fish     |            | money                     |
|     |    | ‘The man will buy fish with the money.’   |    |           |      |          |            |                           |
|     | e. | <b>i-bi-bili</b>                          | ng | lalake    | ng   | isda     | <b>ang</b> | <b>bata</b>               |
|     |    | VC-IRR-buy                                |    | CORE man  | CORE | fish     |            | child                     |
|     |    | ‘The man will buy fish for the child.’    |    |           |      |          |            |                           |

(Foley 2008: 23)

The root of the verb used in the clauses (9a-e) is *bili* ‘buy’. Depending on which argument is preceded by the phrase marker *ang*, one of the following affixes attaches to the root of the

---

<sup>10</sup> Foley (2008: 23) has not included *ang* before *isda* ‘fish’ in his b. example. Because this is most likely an error, I have chosen to include it here.

verb: *-um-*, *-in*, *-an*, *ipaN-* and *i-*, signalling the voice which corresponds to the semantic role of that argument. In (9a), for example, *ang* precedes the actor argument *lalake* ‘man’ and the verb takes the infix *-um-*, signalling actor voice, while in (9c) *ang* precedes the locative argument *tindahan* ‘store’ and the verb takes the suffix *-an* signalling location voice.

The function of *ang* is to mark a particular participant in the sentence as having a special syntactic status. It has been widely debated whether this argument should be referred to as a *subject*. Schachter (1976) argues that the properties typically associated with a subject are divided into two categories: reference-related properties and role-related properties. While both these categories coincide with the subject, in languages that have a typical subject relation, only reference-related properties are associated with the argument preceded by *ang* in Tagalog. These are relativizability, the ability to launch floating quantifiers, and nonoccurrence in existential sentences (Schachter 1976: 290). Role-related properties on the other hand, are associated with the semantic role of actor. They concern the addressee of imperatives, the controller of reflexivization, the target of subjectless infinitival complements and a fixed subject position (Schachter 1976: 290-96). In short, both the argument preceded by *ang* and the actor have special syntactic statuses that are typically associated with the subject relation in languages that have a typical subject. Because of this, there has been a lot of debate as to what the argument preceded by *ang* should be called and terms like ‘syntactic pivot’, and ‘subject’ have both been used (Himmelman 2002: 12). Following Næss (2015: 288) however, I will refer to this argument as the voice-selected argument or VSA.

In short, *ang* marks an argument as voice-selected and the affix that corresponds to the semantic role of that argument attaches to the root of the verb, signalling which voice the clause has. Hence, the examples in (9a-e) illustrate actor voice, patient voice, location voice, instrumental voice and benefactive voice respectively.

This kind of voice system differs from asymmetrical voice systems like the active-passive system in two main ways. First, all of the clauses are morphologically marked for voice in the Philippine-type system. In a typical active-passive system, however, only the passive is morphosyntactically marked for voice, while the active remains unmarked. This is illustrated in the active clause in (10a) and its passive counterpart in (10b).

- (10) a. The bear ate the fish  
 b. The fish was eaten (by the bear)

In the active clause (10a) the verb is unmarked for voice. In (10b) however, the passive is indicated by the auxiliary verb *was* and past participle *eaten*. In other words, the passive clause is derived from the active clause, which is the basic clause in this kind of asymmetrical voice system. In Philippine-type voice systems by contrast, none of the clause types are considered the basic one from which all the others are derived.

Second, only the active form in the active-passive system is transitive, while the passive is inherently intransitive as it only contains one core argument. In (10b), the actor *the bear* is syntactically demoted, resulting in a decrease in valency. In a Philippine-type voice system on the other hand, clauses can alternate between which arguments is the voice-selected one, i.e. is preceded by *ang*, without involving a decrease in valency. Hence, each clause in a Philippine-type voice system is typically considered to be equally transitive.<sup>11</sup>

### 1.3.2 Indonesian-type voice system

The Indonesian-type voice system is found in languages spoken in parts of Sulawesi and the major islands in western Indonesia, including Java, Sumatra, Bali, and Lombok, as well as parts of mainland Southeast Asia and Borneo (Chen & McDonnell 2019: 176). Unlike the Philippine-type voice system, languages with an Indonesian-type voice system only have two voices: actor voice and undergoer voice. Actor voice is typically signalled by a nasal prefix as in the Madurese clause in (11).

(11) Madurese

|                                   |           |        |
|-----------------------------------|-----------|--------|
| Ale'                              | n-(t)oro' | Ebu    |
| younger.sibling                   | AV-follow | mother |
| 'Little brother followed Mother.' |           |        |

(Chen & McDonnell 2019: 175)

The nasal prefix is typically represented by *N-* and is homorganic to the place of articulation of the first consonant of the root (Chen & McDonnell 2019: 178). Depending on the language, undergoer voice is signalled either by a prefix as in the Madurese clause in (12) or by remaining unmarked for voice as in the Balinese clause in (13). The system found in Balinese differs from an active-passive system in that the undergoer voice is unmarked, unlike the

---

<sup>11</sup> Though see Ross (2002: 24-32) and Foley (2008) for discussions on transitivity in Philippine-type languages.

passive, which is typically the morphosyntactically marked option. Furthermore, unlike an active-passive system, both clause types are transitive.

(12) Madurese

|                                   |           |                 |                              |
|-----------------------------------|-----------|-----------------|------------------------------|
| Ebu                               | e-toro'   | Ale'.           |                              |
| mother                            | UV-follow | younger.sibling |                              |
| 'Little brother followed Mother.' |           |                 | (Chen & McDonnell 2019: 175) |

(13) Balinese

|                     |        |        |                              |
|---------------------|--------|--------|------------------------------|
| Bawi-ne             | punika | tumbas | tiang                        |
| pig-DEF             | that   | UV.buy | I                            |
| 'I bought the pig.' |        |        | (Chen & McDonnell 2019: 179) |

The voice-selected argument in (11) is the actor *ale* 'younger sibling' while in (12) and (13) it is the undergoers *ebu* 'mother' and *bawi-ne* 'the pig' respectively. Notice that, unlike languages with a Philippine-type voice system, the voice-selected argument occurs in preverbal position. This is also the case for voice-selected NP-arguments in Äiwoo, as shown in section 1.4 below.

In addition to actor voice and undergoer voice, other semantic relations like benefactives, locatives, and instrumentals can become voice-selected arguments through applicative suffixes which combine with the voice prefixes (Chen & McDonnell 2019: 179). This is illustrated in the Javanese clauses in (14).

- Javanese
- (14) a.    Aku masak        jajan kanggó        Karolina.  
           1SG AV.cook        cake for            Karolina  
           'I baked a cake for Karolina.'                                    (Hemmings 2013: 168)
- b.    Aku masak-aké                jajan Karolina.  
           1SG AV.cook-APPL        cake Karolina  
           'I baked a cake for Karolina.'                                    (Hemmings 2013: 168)
- c.    Karolina        di-masak-aké        jajan  
           Karolina        UV-cook-APPL        cake

‘Karolina was baked a cake.’

(Hemmings 2013: 169)

As seen in clauses (14a-c), the promotion of a benefactive participant involves a two-step process. In (14a), *Karolina* is an oblique argument outside the clause core. The core arguments are *aku* ‘I’ and *jajan* ‘cake’. In (14b), the applicative suffix *aké* promotes *Karolina* into the clause core. Finally, in (14c), *Karolina* is further promoted to voice-selected argument (Hemmings 2013: 168-69).

Finally, languages with an Indonesian-type voice system typically have a construction that is analysed as a passive by Chen & McDonnell (2019: 179). This construction is signalled by a prefix, which in some languages is identical to the prefix signalling undergoer voice, while in others, it is distinct from it (Chen & McDonnell 2019: 179). In the languages where these prefixes are identical, the passive construction is distinguished from the undergoer voice construction depending on the realization of the actor argument. In the passive construction, the actor is either unexpressed or expressed by a prepositional phrase. In the UV-construction by contrast, the actor is typically expressed by an enclitic on the verb. The Standard Indonesian clauses in (15) illustrate a passive construction and a UV-construction respectively.

Standard Indonesian

(15) a. Buku itu di-baca oleh Amir  
book that di-read by Amir  
‘The book was read by Amir.’

b. Buku itu di-baca-nya  
book that di-read-3  
‘The book, (s)he read.’

(Arka & Manning 2008: 47)

In (15a) the actor *Amir* is expressed by the preposition *oleh* ‘by’, while in (15b) the actor is expressed by the enclitic *-nya*.

To sum up, symmetrical voice systems are characterized by alternations between two or more equally transitive constructions, none of which are considered to be the basic one from which the others are derived, as opposed to the *active* construction in active-passive systems. Furthermore, all constructions in a symmetrical voice system are equally marked for

voice (with the exception of the undergoer voice in certain Indonesian-type languages). In an active-passive system on the other hand, the active is typically the default and sole transitive construction and is furthermore unmarked for voice, while the passive is inherently intransitive, derived from the active and is morphosyntactically marked for voice.

Symmetrical voice systems are divided into two categories: Philippine-type languages and Indonesian-type languages. While Philippine-type languages typically have a four-way voice distinction, Indonesian-type languages typically have two, in addition to the applicative constructions. Now that the nature of symmetrical voice systems has been described, I will describe the symmetrical voice system of Äiwoo.

## **1.4 Symmetrical voice in Äiwoo**

In this section, I present the Äiwoo language and its clause system, which is the topic of this thesis. In 1.4.1, I present some general facts about the language before I discuss the main differences between actor voice and undergoer voice in 1.4.2. In 1.4.3, I give a brief description of Äiwoo's third voice construction, namely circumstantial voice. Finally, in 1.4.4, I present the aspect system of Äiwoo, which will be of relevance later in the thesis.

### **1.4.1 The Äiwoo language**

Äiwoo is an Oceanic language spoken by approximately 8 000 people in the Main Reef Islands, as well as on Santa Cruz Island, in Temotu Province of the Solomon Islands in the Pacific Ocean (Næss 2015: 273). Ross & Næss (2007: 470-73) has classified Äiwoo as being part of the Temotu first-order subgroup of the Oceanic branch of the Austronesian language family. This subgroup is further divided into the Reefs-Santa Cruz subgroup, of which Äiwoo belongs, and the Utupua-Vanikoro subgroup, named after the other major islands of Temotu Province. See Figure 1.2 for the geographical location of the Temotu Province, where these languages are spoken.



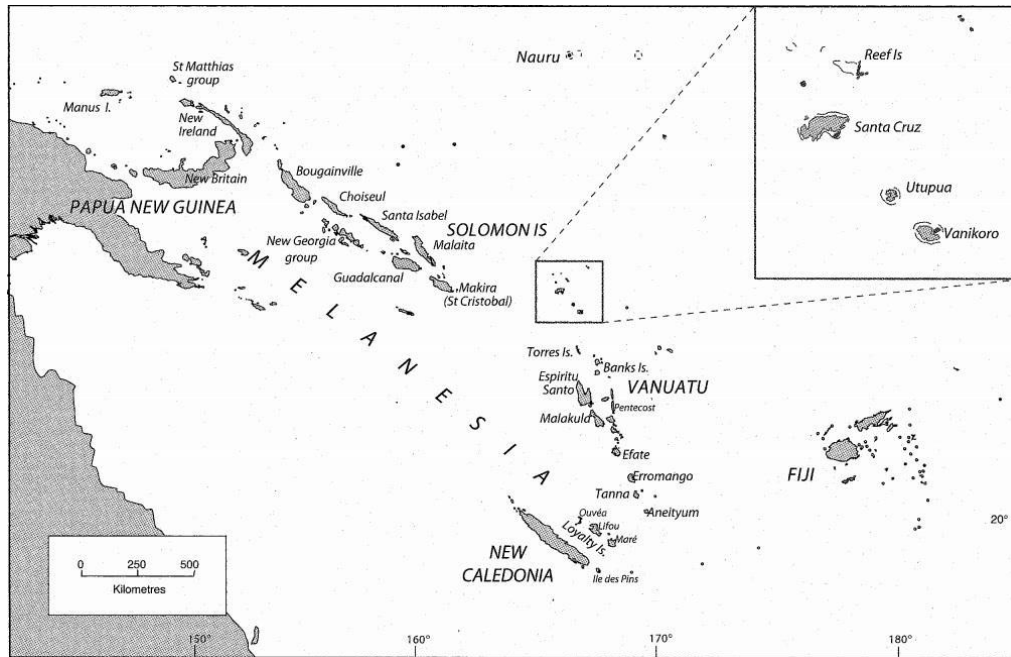


Figure 1.2: Geographical location of the Temotu Province (Ross & Næss 2007: 457)

The data used in this thesis comes from fieldwork conducted by Åshild Næss, and consists of narratives and conversations, which have been transcribed, glossed, and annotated, in the software Fieldwork Language Explorer (FLEX). This corpus consists of 74 different texts of varying lengths and genres. Some of the texts are narratives where the speakers tell different folk tales, children stories and stories from their own lives. Other texts concern different procedures and customs of the reef islanders, where the speakers explain cooking, fishing, hunting, education, marriage, burials, among other aspects of their society. All texts have their own name and each clause is numbered. Hence, *Fishing cairn 314* in (18) below refers to the three hundred and fourteenth clause of the text with the name *Fishing cairn*.

#### 1.4.2 Actor voice and undergoer voice

I now turn to the symmetrical voice system of Äiwoo. Unlike Tagalog and other Philippine-type languages, Äiwoo has a voice system consisting of three voices. These are actor voice, undergoer voice, and circumstantial voice. The primary alternation is between actor voice and undergoer voice. As will become evident below, this is because these voices are expressed through different morphological forms of the verb stem. Circumstantial voice however, is expressed by an enclitic that attaches either to intransitive verbs or to transitive verbs of either voice. Consequently, circumstantial voice must be regarded as secondary to the primary

alternation between actor voice and undergoer voice. This is the reason for investigating the motivation for choice between actor voice and undergoer voice only. In order to provide a complete picture of the voice system in Äiwoo however, I will give a brief description of circumstantial voice in section 1.4.3. For now, I will discuss the main morphosyntactic differences between actor voice and undergoer voice.

Actor voice and undergoer voice are distinguished by the morphological form of the verb. There are different sets of morphological alternations which Roversi (2019: 29-30) analyses as inflectional classes. Table 1.1 lists the different classes.

| Class | Actor Voice | Undergoer Voice | Example  |
|-------|-------------|-----------------|--|
| 1a    | -e          | -i              | <i>läke, läki</i> ‘chop’                       |
| 1b    | -ei/oi      | -i              | <i>gei, gi</i> ‘move, turn’                    |
| 1c    | -Ø          | -i              | <i>lotâlâ, lotälâi</i> ‘prepare’               |
| 2a    | <o> (-u)    | <Ø> (-u)        | <i>tou, tu</i> ‘bring’                         |
| 2b    | <âw> (-ââ)  | <Ø> (-ââ)       | <i>lâwââ, lââ</i> ‘build’                      |
| 2c    | <ow> (-e)   | <Ø> (-u)        | <i>tâlowe, tâlu</i> ‘cut long flexible object’ |
| 3     | (-ei)       | (-i)li          | <i>kei, kili</i> ‘dig’                         |

Table 1.1: Inflectional class of actor voice and undergoer voice (Roversi 2019)

The rightmost column illustrates examples of these patterns. The root of the verb translated as ‘chop’ is *läke* in actor voice and *läki* in undergoer voice.<sup>12</sup> In short, the main voice distinction in Äiwoo is signalled by the morphological form of the verb.

Aside from this main difference, actor and undergoer voice also differ in two other ways, namely in word order and whether the actor argument is expressed by a prefix or a suffix. In actor voice, the actor argument precedes the verb and is marked by a prefix on the verb stem while the undergoer follows the verb (Næss 2015: 273), as illustrated in (16), repeated from (1).

<sup>12</sup> There is no stem *läk-* that can occur by itself – it is always inflected for voice.



In short, in actor voice the actor is marked by a prefix on the verb and occurs in preverbal position. In undergoer voice, by contrast, the actor is marked by a suffix on the verb or is expressed by a postverbal NP, while the undergoer occurs in preverbal position if expressed by an NP, or is postverbal if it is pronominal. Hence, like in Indonesian-type languages, the voice-selected NP-argument typically occurs in preverbal position. While the actor argument is always marked on the verb by an affix, regardless of voice, the undergoer argument is marked on the verb in only a few cases.<sup>16</sup>

To summarize, actor voice and undergoer voice are distinguished by the morphological form of the verb. Furthermore, the two main voices differ in two main ways. The first is which argument appears in preverbal position: the actor in actor voice and the undergoer in undergoer voice. The second is affixal marking on the verb: actors are marked by prefixes in actor voice and by suffixes in undergoer voice.

### 1.4.3 Circumstantial voice

Finally, I come to the third type of voice construction in Äiwoo, namely circumstantial voice. Because the focus of this thesis is the motivations behind the selection of actor voice and undergoer voice, circumstantial voice is not central to this thesis. However, I will give a brief description of it in order to provide a complete picture of the voice system of Äiwoo. Circumstantial voice is signalled by attaching the enclitic =Cä to the verb phrase of either an intransitive clause as in (19), an AV-clause as in (20), or a UV-clause as in (21) (Næss 2015: 284). Note that the verb of the resulting circumstantial voice clause has an actor suffix, regardless of whether the original clause was actor voice or undergoer voice.

- (19) Ilâ                    lu-pwaselee-le                    opo    nugono,                    lâ  
 DEIC:DIST            3AUG-make.AV-UA                    house    betel.leaf                    DEIC.DIST  
 i-mei-i-le=to=wä=nâ  
 PFV-sleep-3AUG-UA=now=CV=DEIC:DIST  
 ‘Then they built a shelter of betel leaves to sleep in.’

- (20) Ko-kä=nä            ku-tu-mu                    mo    ki-tei-mu=wä                    sii=kâ  
 say-DIR:3=CV    IPFV-take-2MIN                    and    IPFV-fish.AV-2MIN=CV    fish=DEIC.DIST  
 ‘(The snake) said: Take it (=a fishing net) and catch fish with it.’

<sup>16</sup> See Roversi (2019) for a detailed description of undergoer marking on the verb in Äiwoo.

- (21) Lâto i-luwa-kä i-konyipe=nä nyibä=nâ  
 thus PFV-take-DIR:3 PFV-wash.UV=CV eye.3MIN=DEIC:DIST  
 ‘Then he took (the magic leaf) and rubbed his face with it.’ (Næss 2015: 284)

The clitic =*Cä* has a number of different functions, most of which are irrelevant for the present thesis. Concerning voice however, the clitic has an important function, namely that of introducing a peripheral or circumstantial participant into the core of the clause. Furthermore, not only does =*Cä* promote a participant from peripheral to core status, but the argument introduced becomes the voice-selected argument like the actor of actor voice and undergoer of undergoer voice. Evidence that the participant introduced by the clitic is a core argument comes from the fact that it is unmarked, unlike circumstantial adjuncts which are most commonly preceded by a preposition. The difference can be seen when comparing sentence (22) and (23) where *chair* in the former is a core argument and is unmarked while in the latter is an adjunct and is preceded by the locative preposition *ngä*. Furthermore, there is a difference in word order: *chair* is preverbal in (22) when it is the VSA, while occurs after the preposition at the end of the sentence when it is a peripheral argument in (23). In short, the VSA is preverbal.

- (22) Chair=kâ ki-li-tokoli-woli=nä, mo i=nâ  
 chair=DEIC.DIST IPFV-GA-sit-go.down=CV CONJ 3MIN=DEIC.DIST  
 ki-so-li=nä.  
 IPFV-stand-go.down=CV  
 ‘A chair is for sitting on, but he is standing on it.’ (Næss 2015: 285)

- (23) Sigiwâu dä i-pu-maa i-tää-e ngä chair.  
 male some PFV-go-LOC:DIST IPFV-sit-go.up LOC chair.  
 ‘A man came and sat down on a chair.’ (Næss 2015: 285)

The participants most commonly introduced by =*Cä* are locatives and instrumentals as illustrated in (19) and (20) respectively. Other types of participants are also possible however, like temporals, stimuli and inanimate causes (Næss 2015: 284).

#### 1.4.4 Aspect and mood marking in Äiwoo

Aside from the voice alternations, another part of Äiwoo grammar will be important for this thesis, namely that of aspect and mood. The aspect and mood markers constitute a single paradigm in Äiwoo. Concerning aspect, the markers most central to this thesis are the perfective and imperfective prefixes. While perfective aspect is expressed by a prefix *i-* as in (24), imperfective aspect is expressed by a prefix *ki-* as in (25) (Roversi 2019: 18).

(24) isäpelivano=nâ      ba      ki-tou=gu  
wife=DEIC.DIST      NEG      IPFV-give.birth=NEG  
'His wife could not have children.'      (Moliki 4)

(25) i-tou=to      dowâlili  
PFV-give.birth=now      child  
'It [the sea snake] gave birth to a child.'      (Moliki 48)

Äiwoo also has two other aspect markers, the postverbal clitics =jo and =to (Roversi 2019: 22). I will not discuss these further as they are not relevant for the topic of this thesis (but see footnote 14 above).

Concerning mood, Äiwoo has two markers; the irrealis prefix *nâ*<sup>17</sup> as, illustrated in (26) and the apprehensive marker *de*, illustrated in (27). The latter is used to signal that an event is undesirable (Roversi 2019: 18).

(26) go      nâ-bi-no  
so.that      IRR.-bake.UV-1MIN  
'So that I can bake'      (Baking potatoes 19)

(27) nyeena      de      baki-ee=dä      ngâ      nuwopa  
tree      APPR      break.UV-go.up=some      LOC.PREP      house  
'[We were worried that] a tree would fall down on the house.'      (Täpeo 45)

In sum, Äiwoo has two main voice constructions: actor voice and undergoer voice, with the addition of a secondary voice construction: circumstantial voice. Furthermore, the aspect and

---

<sup>17</sup> *Nâ-* has a phonological variant *nä-*, as seen in (19) (Roversi 2019: 14).

mood system of Äiwoo consists of four aspect markers: the perfective prefix *i-*, the imperfective prefix *ki-*, and the postverbal clitics *=jo* and *=to*, as well as the two mood markers: the irrealis prefix *nâ-* and the apprehensive marker *de*.

## 1.5 Symmetrical voice or transitivity alternation?

Finally, it is important to discuss the distinctions between a system of transitivity alternations, which is found in most Oceanic languages, and a symmetrical voice system, which is found in western Austronesian languages and in Äiwoo. By making this distinction, it is easier to analyse the clause system of Äiwoo as either one of these systems. This discussion will show that while the notion of *transitivity* is not an easy one to grasp, transitivity alternations in the morphosyntax are often accompanied by differences in what is called *discourse transitivity*. Discourse transitivity is discussed by Hopper & Thompson (1980) and involves the semantic properties of clauses. These properties contribute to determine the degree of *discourse* or *semantic* transitivity of a clause and include kinesis, aspect, mood, affirmation, punctuality, agency, volitionality, and properties of the undergoer argument like affectedness, individuation and animacy. These parameters will be discussed in more detail in chapter 3. For now, it suffices to say that when two clauses differ in terms of formal or morphosyntactic transitivity, these differences are often correlated with a difference in these semantic properties.

The association between morphosyntactic transitivity and semantic transitivity becomes apparent when Margetts (2008) discusses *transitivity discord*, a phenomenon which is common among Oceanic languages. Transitivity discord refers to clauses that are intransitive from a morphological point of view, but transitive from a syntactic point of view. In other words, the verb receives the same type of morphological marking as an intransitive verb but nevertheless takes an object. The clauses in (28a) and (28b) below illustrate the distinction between a regular transitive clause and a transitive discord clause respectively in the Oceanic language Manam.

- Manam
- (28) a. Bóro e u-rere-tá'-di  
 pig this 1SG.RL-like-THC-TR-3PL.OBJ  
 'I like these pigs.' (Margetts 2008: 34)

- b.        *Deparóbu*      *u-rerére*  
             rice                          1SG.RL-like  
             'I like rice (in general).'  
(Margetts 2008: 34)

In both (28a) and (28b), the subject is cross-referenced on the verb by an affix. However, only in (28a) is the object cross-referenced on the verb, signalling that the clause is transitive. While both clauses have an object argument and are thus transitive from a syntactic point of view, only (28a) is morphologically marked as transitive. Hence, there is a mismatch or *discord* between the morphological and syntactical expression of transitivity in (28b): it is syntactically transitive but morphologically intransitive. Thus, languages like Manam have both canonical transitive clauses, i.e. clauses that are both morphologically marked for transitivity and have two arguments, and discord clauses, which also have two arguments but are morphologically marked as intransitives. Consequently, the speaker will have a choice between these two constructions. How then, do speakers choose which construction to use? It is here that Hopper & Thompson's (1980) notion of discourse transitivity comes into play. According to Margetts (2008: 31), while there is some variation between languages as to whether a canonical transitive or a discord clause is selected, the choice between them is correlated with the degree of individuation of the object, or undergoer, as defined by Hopper & Thompson (1980). In general, highly individuated objects correspond to canonical transitive clauses, while objects with a low degree of individuation correspond to discord clauses. In Manam for example, the canonical transitive is chosen if the object is both specific and higher-animal (i.e. human or domestic animal), while the discord transitive is chosen if the object is nonspecific and non-higher animal (Margetts 2008: 33). This distinction is illustrated in (28) above, where the object in (28a) *bóro e* 'these pigs' is specific and higher-animal, while the object in (28b) *deparóbu* 'rice' is neither specific nor higher-animal.

Constructions like the one in (28b) are similar to, but must be distinguished from another kind of intransitive construction with two arguments, namely noun incorporation. These are constructions where the object noun is incorporated into the verb or verbal complex and is not an independent phrase within the clause as opposed to objects in discord clauses. In the Oceanic language Saliba for example, the incorporated object noun appears between the preverbal subject marker and the root of the verb. In a canonical transitive clause however, the object precedes both the root and the preverbal subject marker. The clauses in (29) illustrate this where in (29a) the incorporated noun *niu* 'coconut' is affixed to the verb root *pulisi*



‘husk’ and is preceded by the preverbal subject particle *se*<sup>18</sup>, while in (29b) the same object precedes both the preverbal subject marker and the verb root. The two constructions are further distinguished by the object suffix *-di*, which serves as a morphological marker of transitivity, like in the Manam example in (28a) above.

- Saliba
- (29) a.   Se    niu-pulisi  
           3PL   coconut-husk  
           ‘They husked coconuts.’                                   (Margetts 2008: 34)
- b.   Niu           se    pulisi-di  
           coconut        3PL   husk-3PL.OBJ  
           ‘They husked the coconuts.’                                   (Margetts 2008: 35)

In short, the incorporated object noun is part of the verb and thus differs from the object of a transitivity discord clause. The discord clause (30a) and the transitive clause (30b) illustrate this distinction. Unlike the incorporated object in (29a), the object *Kaleko* ‘clothes’ in the discord clause in (30a) is independent of the verb and precedes the preverbal subject marker, like the objects of the canonical transitives in (29b) and (30b).

- Saliba
- (30) a.   Kaleko       ka           kai-deuli  
           clothes        1EXCL       KAI-wash  
           ‘We washed the clothes / did the laundry.’                   (Margetts 2008: 35)
- b.   Kaleko       ka           deuli-di  
           clothes        1EXCL       wash-3PL.OBJ  
           ‘We washed the clothes.’                                   (Margetts 2008: 35)

Some Oceanic languages have either noun incorporation or discord clauses, while other languages have both. Based on these observations, Margetts (2008: 43) describes a transitivity cline of four different morphosyntactic constructions, where the construction highest on this

---

<sup>18</sup> Referred to as a subject prefix in Margetts’ (2008: 34) terminology.

cline corresponds to high semantic- or discourse transitivity as defined by Hopper & Thompson (1980), while the construction lowest on this cline corresponds to low discourse transitivity. The cline is given in (31).

(31) Transitivity cline

transitive clauses with accord > transitive clauses with discord > intransitive clauses  
with noun incorporation > intransitive clauses

Based on this cline, is it possible to re-analyse actor voice constructions in Äiwoo as a transitivity discord clause? After all, AV-clauses are morphologically identical to intransitives in Äiwoo in that both constructions mark the actor with a prefix on the verb, as opposed to UV-clauses, which mark it with a suffix. Because AV-clauses are morphologically identical to intransitive clauses, but take two arguments, it could be argued that they really are transitive clauses with discord. Consequently, the three different constructions in Äiwoo would correspond to three of the four constructions in Margetts' transitivity cline as illustrated in (32):

(32) One-argument clause with actor **prefix** = intransitive clause

Two-argument clause with actor **prefix** = transitive clause with discord

Two-argument clause with actor **suffix** = transitive clause with accord

According to this analysis, the clause system of Äiwoo is not a symmetrical voice system at all, but rather a system of transitivity alternations. From this analysis, we would expect discord clauses, i.e. AV-clauses to be less transitive from a semantic point of view than accord clauses, i.e. UV-clauses. In particular, Margetts (2008) show that discord objects in the Oceanic languages investigated are clearly less individuated than accord objects in the same languages according to the semantic transitivity framework developed by Hopper & Thompson (1980). Hence, we would expect the undergoers of AV-clauses to be less individuated than the undergoers of UV-clauses.

However, arguments by Næss (2015) clearly show that this is not the case. Rather, undergoer arguments of AV-clauses can be equally individuated as undergoer arguments of UV-clauses. Consider the examples in (33) and (34) which illustrate an AV-clause and an UV-clause respectively.

- (33) Ngaa ... ilâ ki-lää-kä tepekoulâ eââ  
 CONJ DEIC:DIST IPFV-give.AV-DIR:3 thing DEM:DIST  
 ilâ pe-Nganaa=kâ  
 DEIC:DIST COLL-Nganaa=DEIC:DIST  
 ‘So the ones who gave him those things were the people (=spirits) of Nganaa.’  
 (Næss 2015: 280)

The undergoer argument in (33) is the noun phrase *tepekoulâ eââ* ‘these things’ and is clearly individuated. First, it is definite as signalled by the demonstrative *eââ*. Second, it is referential. ‘These things’ refer to “a large luxurious house and all manner of food crops appearing magically overnight in the bush (Næss 2015: 280-1)”. In other words, the undergoer NP refers to particular entities, rather than referring to ‘things’ in general. Because of the definite and referential properties of this undergoer NP, it is clearly individuated. Now consider the UV-clause in (34).

- (34) Ngamaa numonu eângâ ki-la-mä=kâ  
 if money DEM:DIST IPFV-give.UV-DIR:1=DEIC:DIST  
 ki-tokoli mo nyowää?  
 IPFV-sit with what  
 ‘If he gives us money, what will he do with it?’ (Næss 2015: 281)

Here, the undergoer argument is the NP *numonu eângâ* ‘the money’. As signalled by the demonstrative *eângâ*, the undergoer is definite in this clause as well. Hence, the AV- and UV-clauses in (33) and (34) clearly illustrate that the undergoer argument can be individuated in both voices. Consequently, these clauses show that the undergoers of AV-clauses are not less individuated than those of UV-clauses, as we would expect them to be if they were to be analysed as discord-clauses. Because of this, Næss (2015: 281) argues that the alternation between these clauses is best explained in terms of a voice alternation, rather than in terms of semantic transitivity, as discussed by Hopper & Thompson (1980). Furthermore, Næss (2015: 279) argues that this voice alternation is governed by what she calls the *pragmatic prominence* of the participants. However, these arguments are based on individual examples and not on a statistical analysis. Moreover, pragmatic prominence is not easily defined and a systematic study of various factors concerning this notion is needed in order to confirm or

reject this hypothesis. In chapter 3, I discuss how I have investigated various factors relating to pragmatic prominence as well as factors relating to semantic transitivity in order to provide answers to the question of how voice choice is motivated in Äiwoo.

Although the clause system of Äiwoo is presently best analysed as a symmetrical voice system, certain observations suggest that UV-clauses are in the process of being reanalysed as the standard transitive construction in actual speech in Äiwoo. First, UV-clauses are vastly more frequent than AV-clauses. In languages with a symmetrical voice system, it is common for UV-clauses to be more frequent than AV-clauses. Of the 443 transitive clauses in Quick’s (2005, quoted in Riesberg et al. 2020) study on the Sulawesi language Pendau, 45% were AV-clauses, while 55% were UV-clauses. In Cumming’s (1991, quoted in Riesberg et al. 2020) study on Classic Malay, 27% of 115 transitive clauses were AV-clauses, while 73% were UV-clauses. Finally, Riesberg et al. (2020) found in their study of the Sulawesi language Totoli, that 28% of 962 transitive clauses were AV-clauses, while 72% were UV-clauses. There are, however, counterexamples of this, where AV-clauses are more frequent than UV-clauses, e.g. McDonnell’s (2016, quoted in Riesberg et al. 2020) study on Besemah, where 56% of 899 transitive clauses were AV-clauses, while 44 were UV-clauses, and Pastika’s (1999, quoted in Riesberg et al. 2020) study on Balinese, where 62% of 1851 transitive clauses were AV-clauses, while 38% were UV-clauses.

While it is common in western Austronesian languages for UV-clauses to be more frequent than AV-clauses, none of these languages is even close to showing the same difference between UV-clauses and AV-clauses compared to Äiwoo. In Næss’ dataset, the ratio of AV-clauses vs. UV-clauses is roughly 1 to 7 (of approximately 5570 transitive clauses in Äiwoo, 12% are AV-clauses, while 88% are UV-clauses). Furthermore, there are a number of transitive verbs in Äiwoo with no attested actor voice form. This includes high frequency verbs, like *te* ‘see’ and *kä* ‘say’. Finally, in some cases, AV-verb roots are used in UV-constructions even though corresponding UV-roots of the same verb are attested. An example of this is the AV-verb root *mâtâlâ*, as shown in (35).

- |      |                                 |                                     |            |                   |            |
|------|---------------------------------|-------------------------------------|------------|-------------------|------------|
| (35) | <i>lâto</i>                     | <i>ku-mâtâlâ-kâ-no=ngâ</i>          | <i>ile</i> | <i>paveli kâ</i>  | <i>ile</i> |
|      | then                            | IPFV-prepare.AV-DIR3-1MIN=DEIC.DIST | this       | garden DEIC.DIST  | this       |
|      | ‘then I prepare the garden now’ |                                     |            | (Growing pana 66) |            |

While the verb root *mâtâlâ* is morphologically actor voice, the verb takes an actor suffix, not a prefix, indicating that this is a UV-construction, not an AV-construction. This is unexpected, as there are not only one, but two attested UV-forms of this root, *mâtâlâi* and *mâtâlâive*. An example where *mâtâlâi* is used in a UV-construction is shown in (36).

- (36) ku-mâtâlâi-eopu-no=nge  
 IPFV-prepare.UV-also-1MIN=DEIC.PROX  
 ‘[the other one] I prepared as well’ (Thelma S. 20)

Again, the actor suffix indicates that (36), like (35), is a UV-construction. Because *mâtâlâ* has two attested UV-forms, we would expect the speaker to use one of these in the UV-construction in (35). Another example of an AV-verb root being used in a UV-construction is the AV-verb root *wâno* ‘heat’ in (37).

- (37) ku-wâno-no            sii  
 IPFV-heat.AV-1MIN    fish  
 ‘and I heat the fish’ (Fishing for food 133)

Like *mâtâlâ*, *wâno* has a corresponding UV-form, which is *wânoeâ*. An example of *wânoeâ* used in an UV-construction is shown in (38).

- (38) wânoeâ-no    lâ            idea=kâ  
 heat.UV-1MIN   DEIC.DIST    ?=DEIC.DIST  
 ‘when I have heated them’ (Fishing for food 135)

Because *mâtâlâ* and *wâno* are both morphologically AV-verb roots, we would expect them to only be used in AV-constructions, not in UV-constructions as in (35) and (37).

In short, the fact that UV-clauses are vastly more frequent than AV-clauses, that certain high frequency transitive verbs do not have attested AV-forms, and the fact that UV-constructions can be used with AV-verb roots even though these roots have corresponding UV-verb roots all suggest that UV-clauses are in the process of becoming the default transitive clause type in Äiwoo. Consequently, the clause system of Äiwoo might eventually be reanalysed as a system of transitivity alternations like that found in other Oceanic languages. I return to this discussion in chapter 5.

## 1.6 Scientific motivations

In sum, Äiwoo is an Oceanic language with a symmetrical voice system consisting of three voices: actor voice, undergoer voice and circumstantial voice. Unlike Äiwoo, the clause system of most Oceanic languages is characterized by transitivity alternations rather than alternations of voices. Furthermore, a symmetrical voice system is not reconstructed for POc, as it is for PAn and PMP. The two main voices in Äiwoo are actor voice and undergoer voice, where the frequency of the latter outnumber the frequency of the former by approximately seven times. My research question is: What factors motivate a speaker's choice of one voice construction over another?

An answer to this research question will hopefully contribute to the field of linguistic typology in a number of ways. First, an answer to the question of what governs voice choice in Äiwoo will provide a better understanding of the argument structure and voice system of this language. In other words, it will provide a greater insight into the language itself and how it relates to a typical Oceanic system.

Second, an answer to this question will contribute to a better understanding of the nature of symmetrical voice systems in general. Because a symmetrical voice system has been reconstructed for PAn and PMP, but not for POc, one mystery concerns how Äiwoo, being an Oceanic language, has retained a symmetrical voice system when Oceanic languages typically have accusative alignment and no voice alternations. If an answer to this question is found, a better understanding of the link between the symmetrical voice system of western Austronesian languages and of other Austronesian languages will emerge.

Finally, Næss (2015) puts forth a hypothesis that voice choice in Äiwoo is motivated by what Givón (1983) calls *topic persistence*, which refers to how frequently a given referent is mentioned following its current mention in a text or conversation. By answering the question of what governs voice choice in Äiwoo, I will be able to test this hypothesis.

## 1.7 Outline of thesis

This thesis has the following structure. In chapter 2, section 2.1, I present the three different explanatory models I have employed in this thesis (2.1). These models are *topic continuity* as discussed by Givón (1983, 1994), *discourse transitivity* developed by Hopper & Thompson (1980), and the three different *activation states* as discussed by Chafe (1987). In section 2.2, I discuss how five previous studies have applied these models in order to explain the clause

structure of different western Austronesian languages with symmetrical voice systems. These are Walters (1994) on the Philippine language Cebuano (2.2.1), Wouk (1996) on Spoken Jakartan Indonesian (2.2.2), Wouk (1999) on the Indonesian language Sasak (2.2.3), Quick (2005) on the language Pendau, spoken on Sulawesi (2.2.4), and finally McDonnell (2016) on the Malayic language Besemah (2.2.5). I also present a hypothesis by Næss (2015) that voice in Äiwoo is motivated by what is called the *topic persistence* of the voice-selected argument, in section 2.3. In chapter 3, section 3.1, I present my dataset, which consists of 160 Äiwoo clauses and serves as the basis for my investigation. I also discuss how I have employed the three different frameworks, mentioned above, in order to discover which factors motivate voice in Äiwoo. In chapter 4, I present the results from my investigation (4.1) as well as the results from a logistic regression test I have performed in order to see if the results are statistically significant (4.2). I end the chapter with a discussion on independence of observations (4.3), which is a requirement that must be met when performing significance testing. In chapter 5, section 5.1, I summarize the results found in this study, before I, in section 5.2, attempt to draw a conclusion about what governs voice choice in Äiwoo. In section 5.3 I discuss the differences between a symmetrical voice system and a transitivity-based system in light of the results found in this study. Finally, I offer some suggestions on future research in section 5.4.

## 2 Previous studies on voice choice in western Austronesian languages

In the following chapter, I present the three different frameworks employed in this study and how these have been applied in previous studies on voice in western Austronesian languages. The frameworks are presented in section 2.1. I start with discussing *topic continuity* (Givón 1983, 1994) (2.1.1), before discussing *discourse transitivity* (Hopper & Thompson 1980) (2.1.2), and finally *activation states* (Chafe 1987) (2.1.3). In section 2.2, I discuss the previous studies on voice. These are Walters 1994 on the Philippine language Cebuano (2.2.1), Wouk 1996 on Spoken Jakartan Indonesian (2.2.2), Wouk 1999 on the Indonesian language Sasak (2.2.3), Quick 2005 on Pendau spoken on the island Sulawesi (2.2.4), and finally McDonnell 2016 on the Malayic language Besemah (2.2.5). As will become evident throughout this discussion, there are in particular three different frameworks or explanatory models that have often been used to explain clause structure in languages with a symmetrical voice system.

### 2.1 Frameworks

#### 2.1.1 Topic continuity

The first framework that I applied on my dataset in order to explain voice choice in Äiwoo is that of *topic continuity* as proposed by Givón (1983, 1994). The term *topic* in Givón's sense is not a discrete or atomic entity and must be distinguished from the notion of *topic* in other frameworks (e.g. from topic in the topic/focus distinction in generative grammar) (Givón 1983: 5). Rather, the topic is defined as “the participant *most crucially involved* in the action sequence running through the paragraph” (Givón 1983: 8, emphasis in original). In other words, the topic is a participant or referent that has a certain degree of importance in certain parts of the discourse, what Givón refers to as the *thematic paragraph*, defined as thematic units consisting of chains of clauses each carrying their own propositional content (Givón 1983: 7). The importance or *topicality* can be measured by tracking how continuous the referents are. By doing this, Givón developed two ways of measuring a topic's degree of continuity. These are *referential distance (RD)* and *topic persistence (TP)*. Both of these measurements contribute to determine the degree of continuity of a topic.

Referential distance (RD) tracks the number of clauses since the current topic or



referent was last mentioned within the thematic paragraph. This measurement is determined by counting leftwards in a text until the previous mention of the referent is found. The more recently the referent was mentioned, the more continuous it is. Different adaptations of this framework have used different scales of measurement of RD. Givón (1983: 13) operated with a scale from 1 (corresponding to maximally continuous) to an arbitrary upper limit of 20 (corresponding to discontinuous). Givón (1994: 10) on the other hand, used a different scale. If the referent has been mentioned in the preceding clause, it is assigned the value 1, which corresponds to high topicality. If it has been mentioned within the preceding two or three clauses, the value assigned is  $2/3$ , which corresponds to moderate topicality. Finally, if the referent has not been mentioned within the previous three clauses, the value assigned is  $>3$ , which corresponds to low topicality. In short, low referential distance corresponds to high continuity or topicality of the referent, while high referential distance corresponds to low continuity or topicality.

While referential distance tracks how far back the current referent is last found, topic persistence (TP) tracks the occurrence of that referent in subsequent discourse. Again, different adaptations use different scales. Thus, Givón (1983: 15) counts the number of consecutive mentions of the given referent until it is interrupted by an intervening referent, while Givón (1994: 10-11) counts the number of mentions of that referent within the following 10 clauses. According to Givón (1994: 11), referents that are more continuous tend to be mentioned more than 2 times, while referents that are less continuous tend to be mentioned between 0 and 2 times in 10 clauses.

In sum, two measurements are used to determine the degree of continuity of a referent. While referential distance measures how continuous a referent is in preceding discourse, topic persistence measures its continuity in subsequent discourse. According to Givón (1994: 9-10) the two measurements correspond to two separate cognitive dimensions of topicality. Referential distance determines how cognitively accessible the referent is. Thus, if a referent has a low value for referential distance, i.e. the number of clauses since its previous mention is low, it is more cognitively accessible in the mind of the listener. Topic persistence on the other hand, determines how thematically important the referent is, which means that the more mentions it has in the following clauses, the more thematically important it is.

## 2.1.2 Discourse transitivity

The second framework I have employed in this study is that of *discourse transitivity* as discussed by Hopper & Thompson (1980). Hopper & Thompson argue that the notion of *transitivity* is a matter of degree of *effectiveness* of the situation and how much change the action transfers from one participant to another. Furthermore, transitivity is a phenomenon that involves the entire clause, and not only a property of the verb. The degree of effectiveness can be determined by investigating several components or parameters. Each of these parameters have two values, one corresponding to a high degree of discourse transitivity and one corresponding to a low degree of discourse transitivity. The more parameters have transitive rather than intransitive values in a clause, the higher degree of discourse transitivity that clause can be said to reflect. These parameters are given in (39) (Hopper & Thompson 1980: 252).

| (39)                  | High                            | Low                |
|-----------------------|---------------------------------|--------------------|
| A. PARTICIPANT        | 2 or more participants, A and O | 1 participant      |
| B. KINESIS            | action                          | non-action         |
| C. ASPECT             | telic                           | atelic             |
| D. PUNCTUALITY        | punctual                        | non-punctual       |
| E. VOLITIONALITY      | volitional                      | non-volitional     |
| F. AFFIRMATION        | affirmative                     | negative           |
| G. MODE               | realis                          | irrealis           |
| H. AGENCY             | A high in potency               | A low in potency   |
| I. AFFECTEDNESS OF O  | O totally affected              | O not affected     |
| J. INDIVIDUATION OF O | O highly individuated           | O non-individuated |

All of these parameters contribute to the degree to which an effect is transferred from one participant to another, and consequently to the degree of semantic transitivity of the clause. The more parameters that have transitive rather than intransitive values, the higher the degree of semantic transitivity that clause has (Hopper & Thompson 1980: 253).

Hopper & Thompson illustrate how these parameters manifest in the morphosyntax of different languages. Hungarian for instance, makes a grammatical distinction between clauses with referential/individual O's and those with non-referential O's. While non-referential O's precede the verb as in (40a), referential O's follow the verb, as in (40b):

- Hungarian
- (40) a. Péter újságot olvas.  
 Peter paper reads  
 'Peter is reading a newspaper.'
- b. Péter olvas egy újságot.  
 Peter reads a newspaper  
 'Peter is reading a [specific] newspaper.' (Hopper & Thompson 1980: 258)

While the newspaper as an entity is irrelevant to the context in (6a), it is important to the context in (6b). In other words, in (6b) the O is more individuated than in (6a). Consequently, (6b) is more transitive than (6a), which is signalled by change in word order and the presence of an indefinite article *egy* (Hopper & Thompson 1980: 258).

An example of how aspect is correlated with degree of semantic transitivity comes from Finnish. As mentioned above, perfective aspect corresponds to a high degree of semantic transitivity, while imperfective aspect corresponds to a low degree of semantic transitivity. In Finnish, clauses with individuated O's are interpreted as perfective, while clauses with less individuated O's are interpreted as less transitive. This is illustrated in the distinction between (41a) and (41b).

- Finnish
- (41) a. Liikemies kirjoitti kirjeen valiokunnalle.  
 businessman wrote letter(ACC) committee-to  
 'The businessman wrote a letter to the committee.'

- b. Liikemies kirjoitti kirjettä valiokunnalle.  
 businessman wrote letter(PART) committee-to  
 ‘The businessman was writing a letter to the committee.’  
 (Hopper & Thompson 1980: 271)

In (41a) the O is marked by the accusative case, resulting in a perfective interpretation of the clause. In (41b) on the other hand, the O is marked by the partitive case, resulting in an imperfective interpretation. Consequently, (41a) is more semantically transitive than (41b) (Hopper & Thompson 1980: 271). The Hungarian examples in (40) and the Finnish examples in (41), illustrate how degree of semantic transitivity is manifested in grammar.

To summarize, the degree of discourse transitivity of a clause can be determined by investigating different parameters. If several of the parameters correspond to a high degree of effectiveness, a clause can be said to correspond to a higher degree of discourse transitivity. As becomes evident when discussing previous studies that have adopted this framework below, the difference in degree of discourse transitivity between different voices in a symmetrical voice system can say a lot about how voice choice is motivated in these languages.

### 2.1.3 Activation state

Chafe makes a ternary distinction between active concepts, semi-active concepts, and inactive concepts. An active concept is defined as “one that is currently lit up, a concept in a person’s focus of consciousness (Chafe 1987: 25).” Moreover, an active concept represents given information, that is, information that is not new or unknown in the discourse. As such, the speaker assumes that this concept is also active in the mind of the hearer and regards the information associated with it as given for both interlocutors (Chafe 1987: 26). In other words, active information is what is currently being talked about, the topic of the conversation.

The second activation state is semi-active or accessible and is defined by Chafe (1987: 25) as a concept “that is in a person’s peripheral consciousness, a concept of which a person has a background awareness, but which is not being directly focused on.” A concept becomes semi-active or accessible in one of two ways, according to Chafe. First, a concept becomes semi-active when a previously active concept is not refreshed, and thus fades from focus. This concept does not become immediately inactive but remains semi-actively in the interlocutors’

consciousness. When reintroduced then, it does not represent completely new information, but rather accessible information. Second, a concept is semi-active if it is part of a so-called *schema*, which is a set of interrelated concepts that can be expected to surface in the discourse. For example, a schema *class* is associated with the following expected concepts: teacher, students, classroom, and lecture. These concepts then become semi-active or accessible by virtue of being associated with this schema (Chafe 1987: 29).

Finally, Chafe (1987: 25) defines an inactive concept as “one that is currently in a person’s long term memory, neither focally nor peripherally active.” These are concepts that have not been active or accessible earlier in the discourse but is introduced as new information (Chafe 1987: 31).

## 2.2 Previous studies on voice choice

Now that the three different frameworks have been presented, I will discuss how these have been applied by the five previous studies on voice in western Austronesian languages.

### 2.2.1 Walters 1994 – Cebuano

Walters (1994) investigates the clause structure of the Philippine language Cebuano. Like other Philippine languages, Cebuano has a symmetrical voice system with several different voice constructions. Like in Tagalog described above, the voice-selected argument is signalled by a marker of the form *ang*. Walters investigates three different clause types in Cebuano. One of them he refers to as the actor-focus clause, which I will refer to here as actor voice or AV-clause. The remaining two are both object-focus clauses, which signal two different aspects: volitional aspect and potential aspect. The first is the unmarked object focus clause, which I will refer to here as undergoer voice or UV-clause. The other is expressed by a prefix *ma-* on the object focus clause. Following Walters, I will refer to this as the *ma*-clause (Walters 1994: 134).

Walters does not use the term *symmetrical voice* in his paper. Instead, following Cooreman’s et al. (1984) study on the western Austronesian language Chamorro, he seeks to find discourse based evidence that Cebuano is best analysed as being ergative, or what Wouk (1996) defines as *discourse ergative*. According to Wouk (1996: 363), discourse ergativity is determined by identifying structures that serve the same discourse functions as morphological ergativity in languages that are not morphologically ergative. These structures can be defined

as discourse ergative if they are formally identified as passive constructions but are more frequent in texts than structures that are formally identified as active constructions. An example of discourse ergativity is seen when Cooreman et al. (1984) employs the *topic continuity* concept, described in 2.1.1, in order to compare the relative topicality of actors and undergoers in four different transitive clause types: antipassive clauses, ergative clauses, in-passive clauses and ma-passive clauses. By showing that actors are more continuous than undergoers in what they define as the ergative clause type, and by showing that this clause type is most frequently used to express transitive propositions compared to the other clause types, they claim that Chamorro is indeed a discourse ergative language (Cooreman et al. 1984: 14). Likewise, Walters seeks to find similar evidence in order to analyse Cebuano as discourse ergative. He claims that if the actors of UV-clauses are more continuous than undergoers, the UV-clause should be analysed as active-voice clause and consequently that Cebuano is best analysed as an ergative language. If, on the other hand, undergoers are more continuous than actors in UV-clauses, this clause type is best analysed as a passive-voice type. In this case, the ergative analysis is dismissed (Walters 1994: 133).

The question of whether or not Cebuano should be analysed as an ergative language is of little interest to this thesis. What *is* interesting, however, is comparing the topic continuity of the actors and undergoers in the different clause types, which I refer to here as AV-clauses, UV-clauses and ma-clauses. Following the conventions of Givón (1983), referential distance is measured using a scale with an upper limit of 20 in Walters' study. Topic persistence, on the other hand, measures the number of consecutive mentions without interruption and has no upper limit (Walters 1994: 131). Walters applies these measurements to a total of 40 transitive clauses: 24 UV-clauses, 8 AV-clauses, and 8 ma-clauses.

The results show that in UV-clauses, actors are more continuous than undergoers, according to both measurements. Actors had an average RD of 1.3 while the average RD for undergoers was 5.5. Concerning TP, actors had an average value of 2.7 while undergoers had an average value of 0.5. In AV-clauses, actors are less continuous than they are in UV-clauses with an average RD of 5.4 and an average TP of 1.9. Undergoers however, with an average RD of 13.2 and an average TP of 0.8, are nearly as continuous in AV-clauses as in UV-clauses. Of the two arguments, actors are more continuous than undergoers in AV-clauses, but this difference was not as big in UV-clauses. Finally, in ma-clauses, actors are more continuous than undergoers but even less so than in the AV-clauses. In this clause-type, actors had an average RD of 1.3 and an average TP of 1.3, while undergoers had an average RD of

2.9 and an average TP of 0.4 (Walters 1994: 134). Of all three clauses, this was the one where actors are the least continuous and undergoers the most continuous. These results show that only the continuity of actors plays a role in the choice between actor voice and undergoer voice. If the actor is highly continuous, undergoer voice is chosen, while if the actor is less continuous, actor voice is chosen. In other words, if the actor argument is both cognitively accessible, as a low value for RD indicates, and is more thematically important in the following discourse, as a high value for TP indicates, undergoer voice is chosen. This pattern is somewhat unexpected when compared to the results from the other western Austronesian languages discussed in this chapter.

### **2.2.2 Wouk 1996 – Spoken Jakartan Indonesian**

Wouk (1996) investigates the clause structure of Spoken Jakartan Indonesian (SJI). SJI has a symmetrical voice system with two different transitive constructions. Wouk refers to these as actor-trigger clauses and patient-trigger clauses. I will refer to these here as actor voice or AV-clauses and undergoer voice or UV-clauses respectively. While actor voice is signalled by the prefixes *meN-* and *N-*, undergoer voice is signalled by the prefix *di-* (Wouk 1996: 369).

By applying the *topic continuity* concept, Wouk attempts to find evidence of whether or not SJI can be considered to be discourse ergative, much like Walters does for Cebuano and Cooreman et al. does for Chamorro, as seen in 2.2.1 above. Again, discourse ergativity is of little interest to this thesis. It is however, interesting to see how Wouk applies the topic continuity approach to a western Austronesian language with a symmetrical voice system and what the results show about the continuity of the different arguments in the two voice constructions. Wouk's dataset consists of 88 clauses of each voice construction, which were all coded for both referential distance and topic persistence (Wouk 1996: 377).

Following the conventions of Givón 1983, referential distance was tracked using a scale with a maximal limit of 20 clauses. The results show that in AV-clauses, actors are more continuous than undergoers. Using a chi-square test, this difference was shown to be statistically significant. In UV-clauses, undergoers are more continuous than actors. However, this difference was not statistically significant (Wouk 1996: 379). Concerning topic persistence, the results show that in AV-clauses actors are more continuous than undergoers. Again, a chi-square test proved this difference to be statistically significant. In UV-clauses, undergoers are more continuous than actors. This difference was also shown to be statistically significant.

In short, according to the first measurement, referential distance, the voice-selected argument is more continuous than the other. However, this difference was only statistically significant for AV-clauses. According to the second measurement, topic persistence, the voice-selected argument is again shown to be the more continuous of the two. However, here the difference was statistically significant for both voices. This means that when the actor is more continuous than the undergoer, actor voice is chosen; when the undergoer is more continuous however, undergoer voice is chosen. This means that whatever argument is both more cognitively accessible and more thematically important in subsequent discourse than the other governs the choice of the corresponding voice construction. Hence, unlike the pattern found in Cebuano and Pendau (to be discussed below), the continuity of *both* arguments plays a role in voice choice in SJI.

In the same study, Wouk also employs Hopper & Thompson's *discourse transitivity* framework as discussed in section 1.4. Wouk coded 88 of each clause type for all of Hopper & Thompson's transitivity parameters. The results showed that while most parameters show no correlation with either clause type, the parameters punctuality, mood, and individuation of O proved to have some correlation with voice, so that actor voice correlates with a low degree of semantic transitivity and undergoer voice with a high degree of semantic transitivity (Wouk 1996: 372).

First, a correlation was found between mood and voice, so that the observed frequency of AV-clauses with irrealis mood were higher than the expected frequency, while the observed frequency of UV-clauses with irrealis mood were lower than the expected frequency. In other words, irrealis mood, which reflects a low degree of semantic transitivity, tends to occur more often with AV-clauses than with UV-clauses. A chi-square test proved these differences to be statistically significant (Wouk 1996: 373).

Second, approximately one third of AV-clauses were punctual while two thirds were durative. UV-clauses however, showed an opposite pattern: two thirds were punctual while one third were durative. Again, a chi-square test proved this difference to be statistically significant. Because punctual events reflect a higher degree of semantic transitivity and durative events reflect a lower degree, actor voice once again corresponds to a lower degree of semantic transitivity (Wouk 1996: 373-4).

Finally, of the number of sub-parameters that contributes to the status of individuation of O, the most important one proved to be referentiality. Wouk makes a ternary distinction between non-referential, referential but not identifiable, and identifiable O arguments.



Referential NPs are used to refer to objects that have a continuous identity over time in a narrative. Identifiable NPs are referential and can be identified by the listener by virtue of unique reference, prior mention, or inclusion in a previously invoked frame. Not identifiable NPs are referential but not identifiable, unlike the previous category. Finally, non-referential NPs are neither referential nor identifiable (Wouk 1996: 374-5). The results showed that almost half of the patients in the AV-clauses were non-referential, while only 2% of the patients in the UV-clauses were non-referential. Furthermore, while for AV-clauses the ratio of identifiable to unidentifiable referents were 2:1, it was 6:1 for UV-clauses. Yet again, a chi-square test proved this difference to be statistically significant. This demonstrates once again that UV-clauses correspond to a higher degree of semantic transitivity while AV-clauses correspond to a lower degree of semantic transitivity (Wouk 1996: 375).

In sum, Wouk demonstrated that mood, punctuality and individuation of O all showed that there was a significant correlation between UV-clauses and high degree of semantic transitivity and between AV-clauses and low level of semantic transitivity.

### **2.2.3 Wouk 1999 – Sasak**

Next, Wouk (1999) investigated the clause structure of the Indonesian language Sasak, spoken on the island of Lombok. Like Wouk (1996), she employs Hopper and Thompson's discourse transitivity framework, in order to show that this language is different from the other western Austronesian languages that have been subjected to this type of study. Her hypothesis is that if Sasak were to pattern like other Western Austronesian languages that have been investigated using this approach, then actor-focus (nasal) clauses would be expected to correlate with a low degree of semantic transitivity, while patient-focus (oral) clauses would be expected to correlate with a high degree of semantic transitivity (Wouk 1999: 108). In order for her results to be representative of more than one type of natural speech, Wouk investigates both narratives and nonnarrative speech (Wouk 1999: 104). She presents results from three of the parameters from this framework: aspect, mood, and individuation of O. While her hypothesis that Sasak is different from other western Austronesian languages is not of great interest to this thesis, it is interesting to see how the two clause types correlate with semantic transitivity. What Wouk calls actor-focus or nasal clauses, I will refer to here as actor voice or AV-clauses and what she calls patient-focus or oral clauses, I will refer to as undergoer voice or UV-clauses.

First, Wouk investigates aspect. According to Hopper & Thompson, the imperfective

aspect reflects a low degree of semantic transitivity, while the perfective aspect reflects a high degree. Because Wouk expects AV-clauses to reflect a low degree of semantic transitivity and UV-clauses to reflect a high degree, then the former clauses are expected to correlate with imperfective aspect while the latter are expected to correlate with perfective aspect. The results, however, were not very informative, as Wouk's data included no clauses marked with imperfective aspect and only a few clauses marked with perfective aspect. No correlation between aspect and voice were found among the perfective clauses (Wouk 1999: 108-9).

Next, Wouk investigates mood. Concerning this parameter, irrealis mood reflects a low degree of semantic transitivity, while realis mood reflects a high degree. Therefore, AV-clauses are expected to correlate with irrealis clauses while UV-clauses are expected to correlate with realis ones. The results show that in narrative clauses, UV-clauses were correlated with realis mood, while AV-clauses were correlated with irrealis mood. Hence, UV-clauses reflect a higher degree of semantic transitivity than AV-clauses in Sasak narrative clauses according to this parameter. In nonnarrative clauses however, there were no correlation between mood and voice (Wouk 1999: 109).

Finally, Wouk investigates individuation of O. Clauses with highly individuated undergoer (O) arguments reflect a higher degree of semantic transitivity than clauses where the O is less individuated. One common way to determine the individuation of undergoers is to check if it is referential or not. Therefore, Wouk expects AV-clauses to correlate with nonreferential undergoers and UV-clauses to correlate with referential undergoers (Wouk 1999: 109). The results show that this is in fact the case: AV-clauses are found with nonreferential and unidentifiable undergoers, while UV-clauses occur frequently with identifiable undergoers.<sup>19</sup> However, in other western Austronesian languages like Tagalog, this parameter is a conditioning factor, so that referential undergoers guarantee the choice of

---

<sup>19</sup> Wouk (1999: 109) actually claims the opposite when she writes, "Oral forms [UV-clauses] are found with nonreferential and unidentifiable patients, while nasal forms [AV-clauses] occur quite frequently with identifiable patients." I assume this is an error. In fact, on the following page in her article, she presents the following table, which shows that it is indeed oral forms (UV-clauses) that are found with referential patients (undergoers), while nasal forms (AV-clauses) are found with non-referential patients (undergoers).

Table 12. Patient referential status (Wouk 1999: 110)

|       | Non-referential | Interrogative | 1 <sup>st</sup> mention | Identifiable 1 <sup>st</sup> mention | Later mention | Participant in conversation |
|-------|-----------------|---------------|-------------------------|--------------------------------------|---------------|-----------------------------|
| Oral  | 6               | 16            | 9                       | 13                                   | 76            | 22                          |
| Nasal | 30              | 0             | 35                      | 5                                    | 54            | 8                           |

undergoer voice, while non-referential undergoers guarantee the choice of actor voice. Hence, Sasak differs from these languages in that this is merely a tendency and not a rule (Wouk 1999: 109).

In short, Wouk found that UV-clauses correlate with a high degree of semantic transitivity, while AV-clauses correlate with a low degree of semantic transitivity according to the parameters mood (in narrative clauses) and individuation of O. Consequently, the same correlations between voice and degree of semantic transitivity was found for Sasak, as for SJI in Wouk (1996).

#### **2.2.4 Quick 2005 – Pendau**

Quick (2005) investigates the clause structure of the western Malayo-Polynesian language Pendau, spoken on Sulawesi. Pendau has two different transitive clauses, which are signalled by the prefixes *nong-* and *ni-* respectively. In Quick's terminology, *nong*-clauses represent active voice, because the actor is the pivot (what I am here calling the voice-selected argument), while *ni*-clauses represent inverse voice, because the undergoer is the pivot (Quick 2005: 222). According to the terminology adopted in this thesis, I refer to these as actor voice or AV-clauses and undergoer voice or UV-clauses respectively.

Like Walters (1994) and Wouk (1996), Quick employs the topic continuity concept. When measuring topic continuity, Quick applies only one of Givón's measurements, namely referential distance (RD). He uses the conventions from Givón (1994) and distinguishes between two values: 1-3 (the referent has been mentioned in one of the three preceding clauses) and >3 (the referent has not been mentioned in any of the three preceding clauses). In other words, 1-3 corresponds to high topicality while >3 corresponds to low topicality (Quick 2005: 229).<sup>20</sup> Quick then applies this measurement of referential distance to both actor and undergoer in 746 clauses spread over four different narrative texts in Pendau (Quick 2005: 225).

The results show that in UV-clauses, both actors and undergoers are highly continuous, and actors are even more so than undergoers. In AV-clauses however, actors are highly continuous and always more so than undergoers, which vary in their degree of continuity (Quick 2005: 230). From these observations, Quick (2005: 231) concludes that the

---

<sup>20</sup> In this regard, Quick deviates from Givón's (1994) convention, whose scale makes a ternary distinction between 1, 2-3, and >3.

continuity of actors does not motivate voice choice, but the topicality of undergoers however, does; if the undergoer is highly continuous, undergoer voice is chosen, if it is not, then actor voice is chosen. In other words, if the undergoer is more cognitively accessible and more thematically important than the actor, undergoer voice is chosen. Hence, this pattern is opposite to the one found in Cebuano, discussed above, where the continuity of actors govern voice choice and undergoer voice is selected if the actor, not the undergoer, is highly continuous. Furthermore, unlike the pattern found in SJI, and like the one found in Cebuano, only the continuity of one of the arguments plays a role in voice choice.

### 2.2.5 McDonnell 2016 – Besemah

The final study I will be discussing is one conducted by McDonnell (2016) on the Malayic language Besemah, spoken in the highlands of southwest Sumatra in Indonesia. The voice system of Besemah makes a binary distinction between what McDonnell refers to as agentive voice and patientive voice. In line with the terminology in this thesis, I will refer to these as actor voice and undergoer voice, respectively. McDonnell investigates which factors motivate speakers to choose one of the voices over the other. To this end, he employs two explanatory models: discourse transitivity and the activation framework discussed by Chafe (1987).

In order to determine what triggers the selection of one voice construction over the other, McDonnell investigates a number of different parameters that determine the degree of semantic transitivity in Hopper & Thompson's framework, two of which are of particular interest here. They are *animacy* and *generalizability* and are both sub-parameters within Hopper & Thompson's parameter individuation of O. I will discuss each of these in turn, starting with animacy.

First, McDonnell (2016: 210) investigates whether the animacy of the actor and undergoer plays a role in voice choice. He groups the clauses in his corpus into four parts, according to the animacy of the A and O arguments: A-animate:O-animate, A-animate:O-inanimate, A-inanimate:O-animate, and A-inanimate:O-inanimate. The results show that clauses with animate A's were much more frequent than those with inanimate A's. In particular, clauses with animate A's and inanimate O's were the most frequent. Concerning voice choice, the animacy of A does not appear to be a factor. When O is animate however, undergoer voice is more likely to be selected in Besemah (McDonnell 2016: 211). Hence, only the animacy of the undergoer motivates voice choice in Besemah.

Second, McDonnell applies the parameter generalizability, where A and O arguments

are coded as either particular or general. Particular arguments refer to ‘specific individuals or instances of a group’ (McDonnell 2016: 223), whereas general arguments refer to ‘a class of referents or indiscriminately to any member of that class’ (McDonnell 2016: 223).

The results show that in clauses where both A and O are particular, which is the most frequent clause type, undergoer voice is more likely to be selected in Besemah. In clauses where only one of the referents were particular, the voice corresponding to that argument was selected. Thus, actor voice was selected when A was particular and O was general, and undergoer voice was selected when O was particular and A was general. Clauses where both arguments were general were very infrequent. In these cases, there was no preference for voice (McDonnell 2016: 224). These results suggest that undergoer voice is motivated by the status of the undergoer argument, so that undergoer voice will be selected if the undergoer is particular. Otherwise, actor voice is selected. In other words, the status of the undergoer plays a bigger role in voice choice in Besemah than the status of the actor, according to this parameter.

McDonnell (2016) also employs activation states, developed by Chafe (1987). Because of the difficulty in making a ternary distinction however, McDonnell (2016: 220) chose to limit his investigation to a binary one between what in his terminology is called *given* and *new*, which correspond to active and inactive respectively. He coded all actor and undergoer arguments as either given and new in the hopes of determining whether the activation state of these arguments plays a role in the voice choice in that Besemah. His results show that clauses where both A and O are given are by far the most frequent. In these clauses, undergoer voice is more common in Besemah. Actor voice is more common when A is given and O is new. Finally, undergoer voice is more common when A is new and O is given (McDonnell 2016: 221). Again, these results suggest that undergoer voice in Besemah is motivated by the status of the undergoer, so that, if the undergoer is given, undergoer voice is selected.

In sum, McDonnell employs three parameters that all concern properties of the actor and undergoer arguments: animacy, generalizability and activation state. Two interesting observations can be made from his results. First, it is clearly the properties of the undergoer and not the actor that governs voice choice in Besemah. Two of these properties, animacy and generalizability, concern the individuation of the undergoer, which according to Hopper & Thompson (1980) reflects a high degree of semantic transitivity. Second, it is undergoer voice, not actor voice, which is selected if the undergoer argument is animate, particular or

given. This clearly suggests that undergoer voice in Besemah reflects a higher degree of semantic transitivity.

## 2.2.6 Summary of results

In sum, the nature of the voice and clause systems of western Austronesian languages has been widely discussed. What is common among these studies are the frameworks they apply. The three recurring frameworks discussed in this chapter are topic continuity developed by Givón (1983), discourse transitivity developed by Hopper & Thompson (1980) and activation states developed by Chafe (1987).

Concerning the first of these frameworks, topic continuity, the results differ as to whether the continuity of the actor or undergoer determines voice choice in the languages discussed. Walters (1994) found that only the continuity of actor arguments motivates voice choice in Cebuano. Specifically, he found that the choice of undergoer voice is motivated by highly continuous actors. This pattern is somewhat unexpected, when compared to the results from the other studies discussed here. However, Walters' dataset is very limited, consisting of 24 UV-clauses and only 8 AV-clauses. Quick (2005) however, found the opposite pattern for Pendau. The voice choice of this language is motivated by the continuity of the undergoer argument. In contrast to Cebuano then, undergoer voice is chosen if the undergoer, not the actor is highly continuous. Hence, the studies by Walters and Quick have in common that only one of the arguments motivate voice choice in their respective languages, although they differ as to whether this argument is the actor or the undergoer. This is contrasted by the results from Wouk's (1996) study on Spoken Jakartan Indonesian (SJI). Wouk found that the continuity of both arguments motivates voice choice in SJI, so that actor voice is chosen if the actor is more continuous than the undergoer and vice versa. In short, these studies show that voice choice can be motivated by the continuity of either the actor or the undergoer arguments or by both. Furthermore, in all three studies, both referential distance and topic persistence motivate voice choice in these languages. This means that both the cognitive accessibility and the thematic importance of an argument is important for the choice of voice.

Concerning the second framework, discourse transitivity, Wouk (1996) found a correlation between undergoer voice and a high degree of semantic transitivity and between actor voice and a low degree of semantic transitivity according to the parameters mood, punctuality and individuation of O in SJI. Wouk (1999) found the same correlations in Sasak between degree of semantic transitivity and voice according to the parameters mood and

individuation of O. McDonnell (2016) found some interesting correlations between individuation of O and voice in Besemah. He furthermore found a correlation between voice and the animacy of undergoer arguments. In short, these three studies show that the parameters mood, punctuality, individuation of O, and animacy all contribute to determining the degree of semantic transitivity of a clause and can be used to suggest a correlation between voice and semantic transitivity.

Finally, McDonnell (2016) also found some interesting results concerning the third framework: activation states. By making a binary distinction between given and new arguments, McDonnell showed that undergoer voice is more likely to be selected if the undergoer argument is given, rather than new.

In sum, there is a tendency that various properties of the undergoer governs voice choice in the languages discussed in this chapter. The exception to this tendency is Walters' (1994) study on Cebuano, where the topic continuity of the actor motivates the choice of undergoer voice. In the other studies however, undergoer voice is motivated by various properties of the undergoer. Hence, undergoer voice is more likely to be chosen if the undergoer is cognitively accessible, thematically important, animate, referential, particular and given<sup>21</sup>. According to Hopper & Thompson (1980: 253) animate and referential arguments are more individuated, as are particular arguments in McDonnell's study. It can furthermore be argued that arguments that are more cognitively accessible, thematically important and given are more individuated than arguments that are inaccessible, unimportant or new.<sup>22</sup> Hence, it seems that a general tendency in these languages that the choice of undergoer voice is correlated with the individuation of the undergoer, a parameter that corresponds to a high degree of semantic transitivity. Furthermore, the clause related properties mood and punctuality are also correlated with voice in these languages. Again, it is realis and punctual clauses, which correspond to a high degree of semantic transitivity, that correlate with the choice of undergoer voice. Hence, there is a clear tendency in these languages that undergoer voice reflects a high degree of semantic transitivity, while actor voice reflects a low degree of semantic transitivity. I will return to the discussion of the relationship between voice and transitivity in section 5.3.

---

<sup>21</sup> There are obvious overlaps between cognitive accessibility and givenness. I discuss the relationship between topic continuity and activation state in more detail in section 5.1.4

<sup>22</sup> This is discussed in greater detail in section 5.3

## 2.3 Previous hypotheses on voice choice in Äiwoo

Næss (2015: 279) proposes a hypothesis that voice in Äiwoo is governed by the topic persistence of the voice-selected argument. Hence, this hypothesis states that actor voice is selected if the actor is more continuous than the undergoer and that undergoer voice is selected if the undergoer is more continuous than the actor according to the topic persistence (TP) measurement. In other words, actor voice is chosen if the actor is mentioned more frequently than the undergoer in the following discourse, and vice versa for undergoer voice. This hypothesis is based solely on a few individual observations and has not been subjected to any quantitative or statistical analysis. However, a main goal of this thesis is to carry out such an analysis.

In the following chapter, I give a more detailed description of how I have applied Givón's topicality framework in order to test Næss' hypothesis. I also show how I have adopted the other two frameworks by Hopper & Thompson and Chafe in order to investigate how voice choice is motivated in Äiwoo.



## 3 Data and methodology

In this chapter, I will present the data and methodology I have used in attempting to answer the research questions presented in chapter 1. My dataset of 160 clauses by native speakers of Äiwoo and how these are selected is presented in section 3.1. In section 3.2, I discuss how I have adopted each of the three explanatory models discussed in the previous chapter, *topic continuity* (3.2.1), *discourse transitivity* (3.2.2) and *activation states* (3.2.3), in order to answer my research question. Finally, I discuss some problems concerning the terminology employed by Chafe (1987) in section 3.2.4.

### 3.1 Data

My research question concerns how speakers choose which voice to use: actor voice or undergoer voice. However, not all verbs in Næss' dataset are attested for both voices. In order to investigate how this choice is motivated in Äiwoo, it was necessary to find those verbs for which both voices were attested. I mentioned in chapter 1 that there is a great difference in frequency between AV-clauses and UV-clauses. However, it is difficult to determine the exact number of instances of actor voice and undergoer voice constructions. The reason for this relies on how the data is organized in FLEEx. The search function only allows for the search of verb roots. However, the valency-increasing morphology of Äiwoo is complex, so that a clause with an AV-root can be an undergoer voice construction or a circumstantial voice construction if derived by certain affixes or clitics. Hence, when a search for actor voice is made, FLEEx ignores all valency-increasing morphology and the output given will include constructions with actor voice, undergoer voice and circumstantial voice. Furthermore, it can be difficult to decide whether a verb root is an AV-root or an UV-root. For this reason, it is only possible to provide a rough estimate of the difference in frequency between AV-clauses and UV-clauses. There are approximately 670 AV-clauses and approximately 4900 UV-clauses in Næss' corpus. This means that UV-clauses are about seven times more frequent than AV-clauses.

My own dataset was collected by first making a search for all AV-roots in Næss' corpus and then filtering out all derived undergoer voice constructions and circumstantial voice constructions as well as all those AV-clauses with roots that did not have corresponding UV-roots attested. Naturally, this left me with a somewhat limited dataset. Alternatively, I

could have investigated every AV-root, regardless of whether that particular root had a corresponding UV-root in the corpus. The fact that a corresponding UV-root is not found in the available corpus does not imply that it does not exist in the language at all. However, as I wanted to be certain that the speaker has a choice between actor voice and undergoer voice, I chose to limit my dataset to only those clauses with roots that are attested for both voices.

Other verbs were also filtered out, for different reasons. First, several AV-verbs occurred in nominalization constructions, like in (42).

- (42) kele ke                    dee nye-li-ebi-na  
 here DEIC.PROX this NMLZ-GA-bake-NMLZ  
 ‘this is how to bake.’ (Baking potatoes 5)

The Äiwoo verb translated to ‘bake’ *ebi* occurs in its AV-form (the corresponding UV-form is *bi*). However, both the prefix *nye-* and the suffix *-na* signal this verb as a nominalization.

Although Äiwoo nominals often behave like predicates, this nominalization does not behave like a transitive verb, as it takes no actor nor undergoer whose continuity is possible to track. Hence, I have filtered out every instance of nominalizations.

Furthermore, my investigation into what motivates voice choice relies on the properties of both the actor and undergoer arguments. Consequently, only clauses where both arguments can be identified are included in my dataset. An example is given in (43).

- (43) Lâto i-päi-woli-kä=nâ                    läge nyigi  
 then PFV-throw-go.down-DIR3=DEIC.DIST skin one  
 ‘And he threw down a banana skin.’ (Rat and turtle 44)

‘He’ in the English translation refers to the rat, which is one of two main characters in this narrative. The rat is the actor of this clause and can easily be identified from the context. Prior to this clause, the rat has climbed to the top of a banana tree and his friend, the turtle, requested that he throws him a banana, as he is not himself able to climb. Instead of throwing a banana, the rat throws down the skin of an already eaten banana. This banana skin is the undergoer of the clause and is also easily identified as it is overtly expressed by the NP *läge nyigi*. Hence, both the actor and the undergoer are easily identified in this clause. However, there were a number of instances of clauses where either or both arguments were difficult to

identify. In many of these cases, the verb seems to have been used intransitively, based on the semantic interpretation. An example is given in (44).

- (44) i-käve-oli                      go                      i-ki-ebi  
1MIN-remove.stones      go.down                      1MIN-IPFV-bake  
'I take the stones out so I can bake.'  
(Baking potatoes 29)

While Äiwoo transitive verbs often omit the undergoer, it is difficult to determine whether the verb 'bake' is used intransitively or takes an unrealized argument 'potatoes'. Without being able to make a decision regarding this, I chose to follow the English translation, where the verb is used intransitively. Hence, cases like this, where the identification of either the actor or undergoer is difficult, was excluded.

This process left me with 160 clauses, half of which are AV-clauses, the other half UV-clauses. These clauses are collected from a total of 23 texts. While 28 clauses are collected from narrative texts, the remaining 132 are collected from texts describing certain procedures or customs. In the following sections, I will describe the three different methodologies or frameworks I have applied to my dataset, starting with Givón's (1983) topic continuity.

## 3.2 Methodological approach

I will now discuss how I have employed the three different explanatory models, starting with topic continuity, before discussing discourse transitivity and activation states in turn.

### 3.2.1 Topic continuity

The actor and undergoer argument of each of the 160 clauses in my dataset were coded for referential distance and topic persistence. Referential distance was measured using a scale from 1-10, where 1 = maximally continuous, and 10 = not continuous. Hence, the first mention of a referent was automatically assigned the value 10. Thus, a low value for RD corresponds to a high degree of continuity while a high value corresponds to a low degree of continuity. The reason for setting the maximal limit at 10, and not 20 as Givón (1983) does, is because 20 proved to be too high: there were no referents that could be tracked all the way to 20 clauses preceding the current mention. When measuring topic persistence, I followed the

conventions of Givón (1994) by counting how many times a referent is mentioned in the 10 following clauses. Thus, a scale between 0 and 10 were used where 0 = no recurring mention and 10 = 10 recurring mentions. Unlike RD then, a low value for TP corresponds to a low degree of continuity while a high value for TP corresponds to a high degree of continuity.

When counting the values for RD and TP, there were certain clauses that had to be skipped. First, only clauses with referential arguments were counted. This excludes clauses like the one in (45):

- (45) lâ                    pevaio-oli-maa                    mo    lâ                    doo=wâ  
 DEIC.DIST    morning-go.down-LOC.DIST    and    DEIC.DIST    like.that=DEIC.DIST  
 ‘and in the morning it was the same again,’                    (Moliki 213)

This clause contains no referents that could be tracked. As such, when counting either the referential distance or topic persistence of a referent, clauses like (45) were skipped, as they did not contain a place for any referent to occur, whether it is the referent being tracked or a potential intervening referent.

Second, clauses that occur in direct speech were omitted, because direct speech signals an interruption from the overall narrative. Hence, any referent that occurs in direct speech cannot be used to track a referent. An example of direct speech is given in (46).

- (46) ji-nâ-wâ=to                    nou    na-mu                    nâ-ngä-ji  
 12MIN-IRR-go=now    banana    POSS.FOOD-2MIN    IRR-eat-12MIN  
 ‘Let’s go and eat your bananas.’                    (Rat and turtle 24)

Although the bananas mentioned here constitute an essential part of the narrative, they occur in direct speech and are thus part of an interruption from the overall narrative. Hence, clauses in direct speech were skipped, when counting.

To summarize, the actor and undergoer argument of each of the 160 clauses in my dataset were counted for both referential distance and topic persistence. While a low value for referential distance means that a referent has recently been mentioned and consequently corresponds to a high degree of continuity, a high value for topic persistence means that the referent has been continuously mentioned throughout the thematic paragraph and consequently corresponds to a high degree of continuity. This was done in an attempt to see if the cognitive accessibility or thematic importance of a referent are motivating factors for

voice choice in Äiwoo. However, as voice choice can be motivated not only by one, but by several contributing factors, I chose not to limit my investigation to only one methodology. In the following two sections, I discuss two alternative explanatory models, namely *discourse transitivity* and *activation states*.

### 3.2.2 Discourse transitivity

I will now discuss my own way of applying Hopper & Thompsons's transitivity approach in order to investigate whether the discourse transitivity of actor voice and undergoer voice can explain voice choice. I discuss each parameter in turn and how to apply them. Because not all of these parameters are morphosyntactically marked in Äiwoo, there were some difficulties in determining how to code the participants and events. As such, I had to base my decisions on semantic criteria.

**Participants:** 2 or more participants are necessary in order for a transfer of effect to take place. Thus, 2 or more participants reflect a high degree of semantic transitivity while only 1 participant reflects a low degree of semantic transitivity (Hopper & Thompson 1980: 252). My data sample includes only clauses with two or more participants. Consequently, every clause contains both an A argument and an O argument. Hence, I have chosen to omit this parameter from my investigation.

**Affirmation:** This parameter determines whether the event is affirmative or negative. The former reflects a high degree of semantic transitivity while the latter reflect a low degree of semantic transitivity (Hopper & Thompson 1980: 252). My dataset consists of texts describing certain procedures and customs. This means that most clauses describe how to do something, or how a particular event occurs or have occurred. As such, there are very few, if any, clauses that are not affirmative. Hence, I exclude this parameter from my investigation.

**Mood:** As the name suggests, this parameter determines the mood of the clause. An event which is asserted to have occurred in the real world is more effective than an action which is either hypothetical or which did not occur. Thus, realis events reflect a high degree of semantic transitivity while irrealis events reflect a low degree of semantic transitivity (Hopper & Thompson 1980: 252). Äiwoo makes a distinction between realis and irrealis. However, there were only six cases of irrealis verbs. Because of this, it is difficult to determine whether mood could be a motivating factor for voice choice. I have therefore chosen to exclude this parameter.

**Aspect:** This parameter makes a distinction between telic and atelic events. While the

former views the event from its endpoint, where the transfer is complete, the latter views the event as ongoing, where the transfer is only partially complete. Thus, telic events reflect a high degree of semantic transitivity while atelic events reflect a low degree of semantic transitivity (Hopper & Thompson 1980: 252). As explained in section 1.4.4, Äiwoo realis clauses make a grammatical distinction between perfective aspect, signalled by the prefix *i-*, and imperfective aspect, signalled by the prefix *ki-*. I have coded all the realis clauses for aspect in order to see if there is a correlation between aspect and voice. If one of the voice constructions correlates with perfective aspect, while the other correlates with imperfective aspect, then the former will reflect a higher degree of semantic transitivity than the other one.

**Kinesis:** Along with the following three parameters, kinesis is not something that is grammatically coded in the language. Therefore, I have used semantic criteria when coding clauses according to this parameter. Hopper & Thompson (1980: 252) make a distinction between actions and states and argue that actions denote more semantically transitive events than states because the former can be transferred from one participant to another while the latter cannot. Following Hopper & Thompson then, I have chosen to code each clause as either action or state. Hence, if one of the voices correlates with actions, that voice will reflect a higher degree of semantic transitivity, according to this parameter.

**Punctuality:** Punctual actions, whose effects are more immediate, correspond to a higher degree of transfer from one participant to another, than non-punctual actions, whose effects are on-going. Thus, punctual actions reflect a high degree of semantic transitivity while non-punctual or durative actions reflect a low degree of semantic transitivity. This parameter was also difficult to code, as I had to base my decisions solely on semantic grounds. Hopper & Thompson describe punctual events as “actions carried out with no obvious transitional phase between inception and completion” (Hopper & Thompson 1980: 252). Furthermore, Saeed (2015: 116) describes the distinction between a punctual and durative event: “durative is applied to verbs that describe a situation or process which lasts for a period of time, while punctual describes an event that seems so instantaneous that it involves virtually no time.” I have attempted to code each of the clauses in my dataset as either punctual or durative according to this definition. However, I met with several difficulties when attempting to code the clauses. An example of an unambiguous punctual action is given in (47).

- (47) mo    nâ-bwii-eäli-mu  
 and    IRR-spit-clear-2MIN  
 ‘then you spit at him.’ (Fishing 333)

The action of spitting at someone is so brief that one cannot identify any notable period of time between start and completion. Hence, this is a clear example of a punctual action. Other predicates, like *cut*, are more difficult to determine according to this parameter. Take for example the event described in (48).

- (48) mo    li-läke          nyânumowä  
 and    3AUG-cut          nyânumowä tree  
 ‘then they cut a nyânumowä tree.’ (Shark net 42)

I have chosen to code this clause as punctual, as the act of cutting something is more or less instantaneous. However, one might argue that cutting a tree is an ongoing process that takes a given amount of time to complete. Nonetheless, the action is clearly punctual if the tree is cut down with a single strike. Now, compare example (48) with the one given in (49):

- (49) lâ                  li-ange          nâbo          lâ                  lu-mâlu=kâ  
 DEIC.DIST    3AUG-slice    dried.breadfruit    DEIC.DIST    3AUG-stop-DEIC.DIST  
 ‘until they have finished cutting nâbo.’ (Fermenting breadfruit 15)

Whereas (48) describes the action of cutting a single entity, (49) describes an action where several breadfruits are being sliced into pieces. Hence, unlike the action in (48), this action takes place over a longer duration of time. I have therefore chosen to code this clause as durative.

Finally, (50) provides an example of a clause which is unambiguously understood as being durative.

- (50) ki-vili-ngopu=waa          nyige          nenu  
 IPFV-weave-1AUG=FUT          coconut.leaflet    coconut  
 ‘we weave them from coconut leaves.’ (Basket weaving 13)

‘Them’ in this clause refers to baskets that the Reef Islanders use to carry on their heads. Because weaving baskets, among other things, is an action that requires a longer duration of time, I have chosen to code this as durative.

In short, while instantaneous actions, like those in (47) and (48) were coded as punctual, actions whose duration lasts a given amount of time, like those in (49) and (50) were coded as durative. If the results show that one voice correlates with punctual actions while the other with durative actions, the former will reflect a higher degree of semantic transitivity according to this parameter.

**Volitionality:** This parameter determines whether the action carried out is voluntary or involuntary from the part of the actor. A transfer of effect is typically more apparent when carried out with volition than one where the actor is an involuntary instigator or participant. Thus, volitional actions reflect a high degree of semantic transitivity while non-volitional actions reflect a low degree of semantic transitivity (Hopper & Thompson 1980: 252). This is another parameter that is not grammatically manifested in Äiwoo clauses. Thus, I have relied on semantic criteria in order to make decisions about how to code clauses according to this parameter. Specifically, I have looked to the semantics of the actor, as this parameter concerns the volitionality of this argument. I have chosen to code an action as volitional if it is performed or set in motion by the actor’s own volition. Thus, the action described in (11) above, is coded as volitional, because the actor is wilfully and purposefully weaving a basket. An example of a non-volitional action however, is given in (51).

|      |  |           |         |             |
|------|--|-----------|---------|-------------|
| (51) | Lu-mo-le                                 | i-tou     | sigiläi |             |
|      | 3AUG-stay-UA                             | PFV-bring | male    |             |
|      | ‘After a while she gave birth to a boy.’ |           |         | (Dolphin 8) |

While it can be argued that the act of giving birth involves active participation from the actor, she cannot choose to set this action in motion at her own volition. Hence, I have chosen to code this clause as non-volitional. If the results show that one voice construction correlates with volitional actions while the other with non-volitional ones, the former will reflect a higher degree of semantic transitivity according to this parameter.

**Agency:** It was difficult to distinguish this parameter from the previous one, as there are obvious overlaps between the two. Furthermore, Hopper & Thompson (1980) do not





Here the undergoer argument ‘house’ is brought into existence by the action performed by the actor. Hence, the undergoer can be said to be fully affected by the action. Clauses with fully affected arguments like this reflect a high degree of semantic transitivity according to this parameter.

Secondly, partially affected arguments are those that undergo a change in location or possession, but not in form or existence. Hence, this category corresponds to what Beavers (2011: 339) calls “directed motion” predicates. Predicates whose meanings are translated to ‘bring’, ‘lift’, ‘give’, ‘share’, and ‘put’ all have partially affected undergoers. An example is given in (53).

- (53) I-pi-mä=nâ    ngâ      nuwopa  
PFV-bring-DIR1=DEIC.DIST      LOC      house  
‘She brought her to the house.’ (Moliki 54)

The undergoer argument that ‘her’ refers to is baby Moliki, whom her mother has just adopted from the sea snake. The baby is brought to the house of her new parents and thus undergoes a change in both location and possession. However, the baby is not physically affected by the action in such a way that her state or existence is altered. Hence, this argument is coded as partially affected.

Finally, arguments coded as not affected undergo neither changes in form or existence nor in location or possession. Predicates like *see* have undergoer arguments that are not affected at all, as exemplified in (54).

- (54) Ki-amoli-kä-de=ngâ  
IPFV-look.at-DIR3-12MIN-DEIC.DIST  
‘We see it.’ (Hunting turtles 19)

The undergoer argument that ‘it’ refers to is the turtles that the hunters are going to catch. No change, either in state, existence, position or possession, befalls the turtles as a result of being seen by the hunters. Hence, I have chosen to code these arguments as not affected.

In short, while fully affected arguments undergo a change in state or existence, partially affected arguments undergo a change in location and/or possession but not in form or existence. Finally, arguments that are not affected do not undergo change in any form. I have coded each of the clauses in my dataset as either one of these three degrees of affectedness of

O in order to find a correlation between this parameter and voice choice. According to Hopper & Thompson, clauses with undergoers that are fully affected reflect a higher degree of semantic transitivity than those with undergoers that are only partially affected or not affected at all. Hence, if one of the voices correlates with a higher degree of affectedness of O than the others do, that voice will reflect a higher degree of semantic transitivity according to this parameter.

**Individuation of O:** This parameter determines the degree to which the undergoer can be individuated or distinguished from both the actor and from its own background. This parameter consists of a number of sub-parameters that contribute all to determine the degree of individuation. If an undergoer is proper, human/animate, concrete, singular, count and referential/definite it is individuated to a higher degree than if it is common, inanimate, abstract, plural, mass and non-referential. Thus, if the undergoer is individuated to a higher degree according to these sub-parameters, then the action reflects a high degree of semantic transitivity (Hopper & Thompson 1980: 252). Because of the amount of time it would take to code each undergoer argument for each of these sub-parameters in order to determine individuation, I have chosen not to do so. Instead, I have adopted a method of determining the individuation of undergoers, which is applied by McDonnell (2016) and is discussed below.

Previous studies that have applied this parameter have mainly considered referentiality when determining the individuation of O. For instance, Wouk (1996: 374) claims that degree of referentiality proved to be the most important aspect for determining patient status in SJI. Furthermore, according to Wouk (1999: 104), in most of the Western Austronesian languages with a symmetrical voice system, individuation of O as determined by referentiality is the single most important factor in determining voice choice.

McDonnell, on the other hand, uses the term *generalizability*, where he distinguishes between *general* and *particular* arguments. As mentioned above, particular arguments refer to “specific individuals or instances of a group.” In other words, this seems to correspond to *referentiality*. Based on this observation, I have coded undergoer arguments for generalizability according to McDonnell’s definition. Thus, if the results seem to indicate that one of the voice constructions corresponds to particular referents while the other corresponds to general referents, the former will reflect a higher degree of semantic transitivity according to this parameter.

**Animacy/SAP:** Finally, I have coded both actors and undergoers according to their animacy. Following McDonnell, I have made a binary distinction between animate and

inanimate. However, unlike McDonnell, I have also coded arguments according to whether they are speech act participants (SAPs) i.e. first or second person referents, or non-SAP's, i.e. third person referents.

By applying this framework, I want to investigate whether either AV-clauses or UV-clauses correspond to a higher degree of discourse transitivity than the other. In particular, it is interesting to find out if UV-clauses correspond to a higher degree of semantic transitivity, considering that these clauses are vastly more frequent than AV-clauses. If this is shown to be the case, these results might provide evidence that the symmetrical voice system of Äiwoo is in the process of being reanalysed into a system of transitivity alternations like that found in other Oceanic languages. Furthermore, as became evident when discussing previous studies in section 2.2, the properties of the undergoer seem to play a central role in the choice of undergoer voice in various western Austronesian languages. Hence, it is also of particular interest to investigate whether this is also the case in Äiwoo.

### 3.2.3 Activation state

It was often difficult to decide the activation state of different arguments. In particular, it was difficult to decide when an argument should be coded as semi-active or active, or semi-active or inactive. As already mentioned, McDonnell chose to make a binary distinction between given and new because of this difficulty (McDonnell 2016: 220). Nevertheless, I decided to make a ternary distinction between active, semi-active, and inactive. In the following section, I will give an example of each of the activation states and provide justifications for my choices. First, an active concept is illustrated in (55).

- (55) lâ                    ki-vili-woli-ngopu=to                    kinyibä                    kâ  
 DEIC.DIST            IPFV-weave-go.down-1AUG=to    IPFV-baskets            =kâ  
 ‘and we weave them into baskets’                    (Basket weaving 19)

The undergoer argument in this clause is *coconut leaves*, which the speaker has been talking about at length in the preceding discourse. This argument has been the focus of attention in the seven preceding clauses. Because this argument has been the focus of the conversation, it will naturally be active in the mind of both interlocutors. Hence, I have coded this argument as *active*. Incidentally, the referent itself is not overtly expressed, which is very common of given undergoers in Äiwoo. This is in line with how Chafe describes the typical verbalization

of active concepts, namely that they are often pronominalized or omitted, or if mentioned, spoken with an attenuated pronunciation (Chafe 1987: 26).

Second, an example of a semi-active concept is given in (56).

- (56) i-pongä          nyopwä=kâ  
 PFV-heat.oven    oven=DEIC.DIST  
 ‘she lit the fire in the oven.’ (Moliki 209)

The semi-active argument in this clause is the oven. This argument has been mentioned several times earlier in the discourse, and it should be well known to both interlocutors that the main characters in this narrative is in possession of an oven. As such, it does not represent completely unknown. However, this argument has not been mentioned in more than ten preceding clauses. In the meantime, the narrator has been talking about how Moliki, the main character, has gone fishing using her special fishing net. Therefore, this fishing net has been active in the meantime, and the oven has not been mentioned once throughout this part of the discourse. Thus, when the oven is finally mentioned again, it is neither completely active nor completely inactive, but somewhere in between – semi-active. This illustrates one of the ways a concept becomes semi-active, according to Chafe, namely by fading out of the active state by not being refreshed through mention. The other way in which a concept becomes semi-active, is by being part of a schema.

Finally, an example of an inactive concept is illustrated in (57).

- (57) lâto    me-ku-wäsele-kä=näa=kâ    vängä  
 then    1AUG-IPFV-make-DIR3=FUT    meal  
 ‘Then we prepare a meal.’ (Burials 56)

The inactive argument is the undergoer *vängä* ‘a meal’. This argument has not been mentioned earlier in the discourse. Hence, it is coded as inactive.

I have coded the undergoer arguments of each clause in my dataset according to this parameter in order to see if the activation state of undergoer arguments plays any role in voice choice in Äiwoo. As for actor arguments, these were difficult to code for activation state. This is mainly due to the type of texts to which I had access. Most of the texts from which my dataset is collected concern different types of procedures (fishing, cooking food) and customs

(education, burials, marriage). In the clauses of these texts, the majority of actor arguments are generic, and do not refer to specific individuals. Take the sentence in (58) for instance.

(58) bulaape-ute-maa                   mo   lâ                   mi-ki-ei=kâ  
tomorrow-again-LOC.DIST          and   DEIC.DIST          2MIN-IPFV-peel=DEIC.DIST  
'the next day you peel them.'   (Making nâbo 40)

This text concerns the procedure involved in making *nâbo*, which are dried breadfruit chips. The actor argument 'you' is clearly a generic *you* and does not necessarily refer to the hearer. Rather, 'you' in this clause refers to any person who undertakes the procedure of making *nâbo*. Aside from these types of texts, there were a total of 22 clauses of the narrative genre. The actor arguments of the clauses were exclusively coded as active. Because of this, it was not possible to determine whether the activation state of actors plays a motivating role in voice choice in Äiwoo, based on my dataset. Consequently, I have limited my investigation of activation state to undergoer arguments. In short, I have coded the undergoer of every clause according to this parameter in order to see if the activation state of O is a motivating factor for voice choice in Äiwoo.

### 3.2.4 Note on terminology

Before discussing the results of my study, a few words on terminology are necessary. Chafe (1987: 21) admits that terms like *old*, *given* and *new* information are confusing and easily misunderstood, especially because they tend to be used with a variety of different meanings in different frameworks. However, he continues to use terms like *active concepts*, *given information* and *given concepts* interchangeably, which only adds to this confusion. He defines *given* as “already active”, *new* as “previously inactive” and *accessible* as “previously semi-active” (Chafe 1987: 22). Hence, it is difficult to distinguish *givenness* from *activation states*. Indeed, McDonnell (2015: 219-20) makes no attempt to distinguish the two terms when he translates Chafe’s ternary distinction between active, semi-active and inactive into given and new, while omitting the intermediary level semi-active. Furthermore, McDonnell uses the terms *given referent* and *new referents* without taking into account the fact that the activation states apply to concepts and not to referents according to Chafe (1987: 25).

In this thesis, I use the terms *active*, *semi-active* and *inactive* to refer to *types of entities*, i.e. humans, animals or inanimate objects. These types of entities are not specific

referents but concepts that refer to that entity in general. For example, a concept 'fish' will not refer to a specific fish, but any fish in general. Hence, activation is not the same as definiteness. A concept can be active without being definite.

In sum, I have employed these three explanatory models as discussed by Givón (1983, 1994), Hopper & Thompson (1980) and Chafe (1987) in order to find out which factors motivate voice choice in Äiwoo. In the next chapter, I present the results from my study as well as the results of a statistical significance test performed in order to show which of these factors proved to be significant predictors for voice.

## 4 Results and analysis

In this chapter, I present the results from my study. In section 4.1.1, I present the results from Givón's (1983, 1994) topic continuity framework applied to all actors and undergoers in my dataset. In section 4.1.2, I present the results of applying Hopper & Thompson's (1980) transitivity framework. Finally, in section 4.1.3, I present the results from Chafe's (1987) activation state framework. I present the results from a logistic regression test in section 4.2, which show that the only statistically significant predictors for voice is aspect and activation state of O. Finally, I discuss some potential problems with applying a significance test on my dataset in section 4.3. This problem concerns independence of observations and I conclude that the 160 clauses in my dataset are indeed independent of each other.

### 4.1 Results

#### 4.1.1 Topic continuity

The continuity measurements referential distance (RD) and topic persistence (TP) were counted for both the actor and undergoer arguments in each of the 160 clauses. The results are given in the box- and whisker plots below, which illustrate how the data is distributed. The bold line within the box represents the median, which separates the lower 50% of the observations from the upper 50% of the observations. The box itself represents the interquartile range, which is the range from the first to the third quartile. This means that 25% of the observations fall between the median and the upper line of the box, i.e. the third quartile, and 25% of the observations fall between the median and the lower line of the box, i.e. the first quartile. Consequently, 50% of the observations are found within the box, while the remaining 50% fall outside the box. The dotted vertical lines are the whiskers, which extend to maximally 1.5 times the interquartile range (Baayen 2008: 30). Each point falling outside the horizontal lines at the end of the whiskers (as seen in figure 4.3 and 4.4) are outliers.



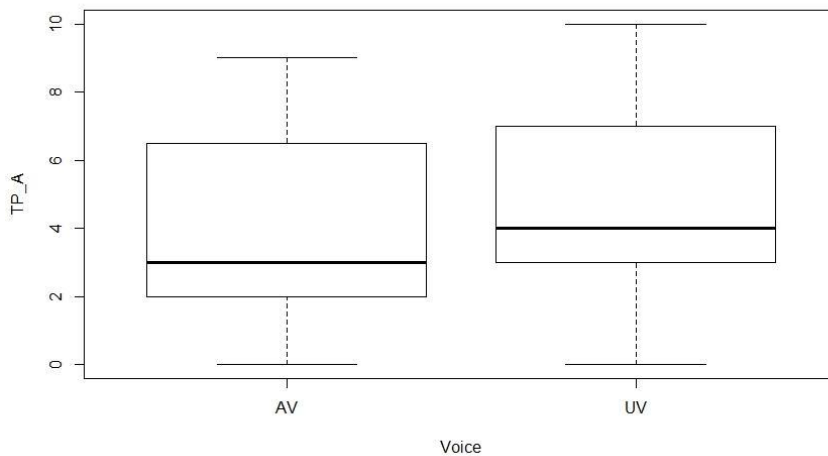


Figure 4.1: Topic persistence of A

Figure 4.1 shows the values for topic persistence of actors in the AV- and UV-clauses in my dataset. As seen from the vertical lines at the top and bottom of the whiskers, the value in actor voice range from 0 to 9, while in undergoer voice it ranges from 0 to 10. The median value in actor voice is 3, meaning 50% of the actors have a higher value than 3 and 50% have a lower value than 3. In undergoer voice, the median is 4 in undergoer voice. This means that while the actors are not particularly low or high in either voice, there is a slight difference in the distribution of the TP-value for actors between actor voice and undergoer voice. As shown below however, this difference is not statistically significant. Consequently, the topic persistence of actors does not seem to be a motivating factor for voice choice in Äiwoo, based on these results.

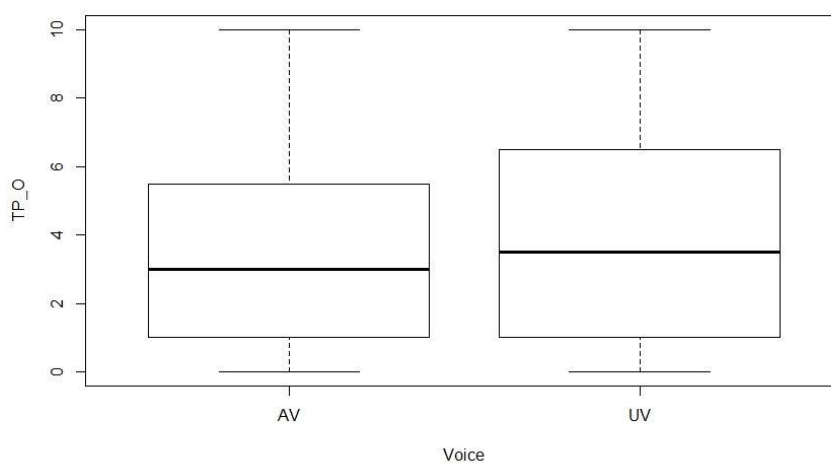


Figure 4.2: Topic persistence of O

Figure 4.2 show the values for topic persistence of undergoers. The values range from 0 to 9, in both AV- and UV-clauses, as seen from the vertical lines at the end of the whiskers. The median value in AV-clauses is 3, meaning 50% of the undergoers have a higher value than 3 and 50% have a lower value than 3. In UV-clauses, the median is 3.5 in undergoer voice. Again, there is a slight difference in the distribution of TP value for undergoers in actor voice and undergoer voice. However, this difference is also shown not to be statistically significant. Like the topic persistence of actors, the topic persistence of undergoers does not seem to be a motivating factor for voice choice in Äiwoo, based on these data. Contrary to Næss' (2015) hypothesis then, my data indicate that the topic persistence, or thematic importance, of the voice-selected argument does not govern voice in Äiwoo.

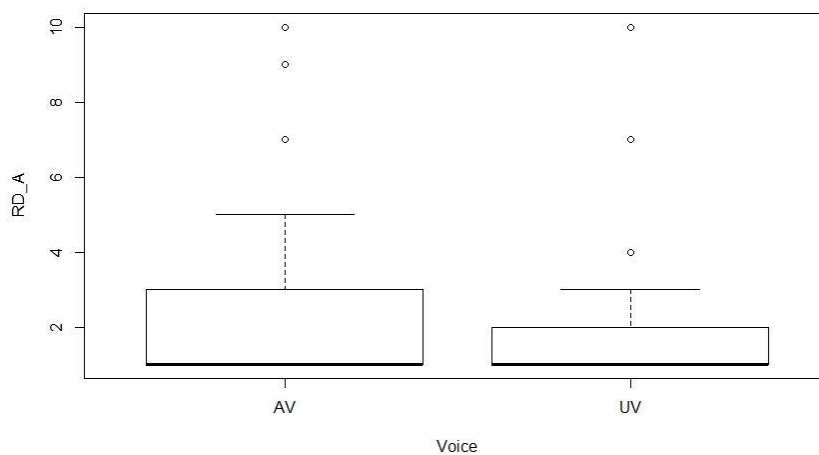


Figure 4.3: Referential distance of A

Figure 4.3 shows the values for referential distance of actors in the AV- and UV-clauses in my dataset. Because the median is so low in both clause types, approximately 1 in value, there is a right, or positive, skew in both distributions. This means that half of the observations have a value of 1. In the AV-clauses, the third percentile is at 3, meaning 75% of the actors have a referential distance between 1 and 3. In the UV-clauses, the third percentile is at 2, meaning 75% of the actors have a referential distance of either 1 or 2. Consequently, the referential distance of actors are very low in both AV- and UV-clauses, indicating that actors are highly continuous, and reflect a high degree of cognitive accessibility, regardless of voice. While there is a slight difference between actor voice and undergoer voice, this difference is not proved to be statistically significant, as discussed below. Yet again, these data indicate that this factor is not a motivating factor for voice choice in Äiwoo.

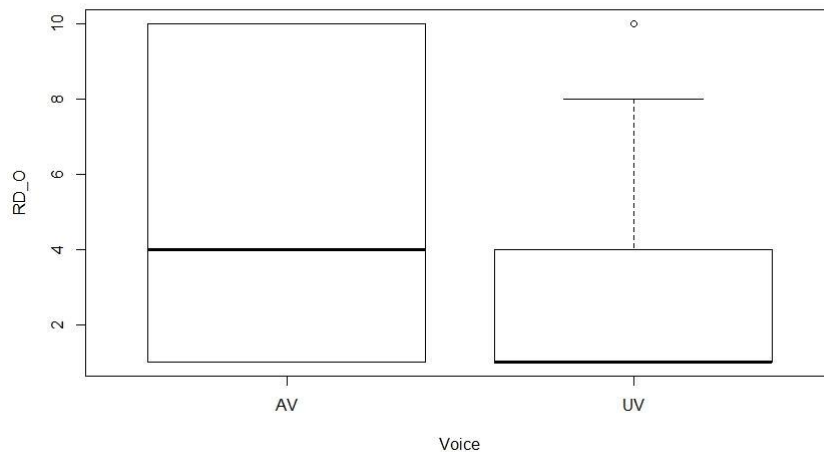


Figure 4.4: Referential distance of O

Finally, figure 4.4 shows the values for referential distance of the undergoers. In AV-clauses, these values vary greatly, with 50% having a value of 4 or lower, and the other 50% having a value of 4 or higher. In UV-clauses on the other hand, 75% of the undergoers have a referential distance of 4 or lower and 50% of them have a value of 1. This means that undergoers are highly continuous in undergoer voice. Furthermore, there is a notable difference in referential distance of undergoers in actor voice and undergoer voice, indicating that the referential distance of undergoers might be a motivating factor for voice choice in Äiwoo. However, yet again this difference proved not to be statistically significant as discussed in section 4.2.

### 4.1.2 Discourse transitivity

In this section, I present the data after annotating all 160 clauses in my dataset according to those of Hopper & Thompson's transitivity parameters that I discussed in the previous chapter. I discuss in turn kinesis, punctuality, volitionality, affectedness of O, individuation of O and finally animacy and SAP vs. third person referents. As will become apparent, the only parameters that differ in the least between actor voice and undergoer voice is aspect, affectedness of O and individuation of O.

## Kinesis

The first parameter I investigated was kinesis, which has two values depending on whether the predicate denotes an action or not (non-action). The results show that almost all of the 160 clauses were coded as action. Only three were coded as non-action. All of these occurred in UV-clauses and had the predicate *eâmoli* “see”, like in (59).

- (59) ki-amoli-kä-de=ngâ  
IPFV-look.at-DIR3-12MIN=DEIC.DIST  
‘we see it.’ (Hunting turtles 19)

As discussed in section 2.2, predicates like this are coded as non-action.

Because only three clauses are coded as non-action, there is not enough evidence in favour of any correlation between kinesis and voice in Äiwoo.

## Aspect

The next parameter I investigated was aspect, which showed far more interesting results than the previous one did. Of the 160 clauses in the corpus, 130 were aspectually marked as either perfective or imperfective aspect. Of these 130 clauses, 62 were AV-clauses, while 68 were UV-clauses. Of the AV-clauses, 12 (19.35%) were perfective, while 50 (80.65%) were imperfective. Of the UV-clauses, 30 (44.12%) were perfective, while 38 (55.88%) were imperfective. These data are illustrated in figure 4.5.

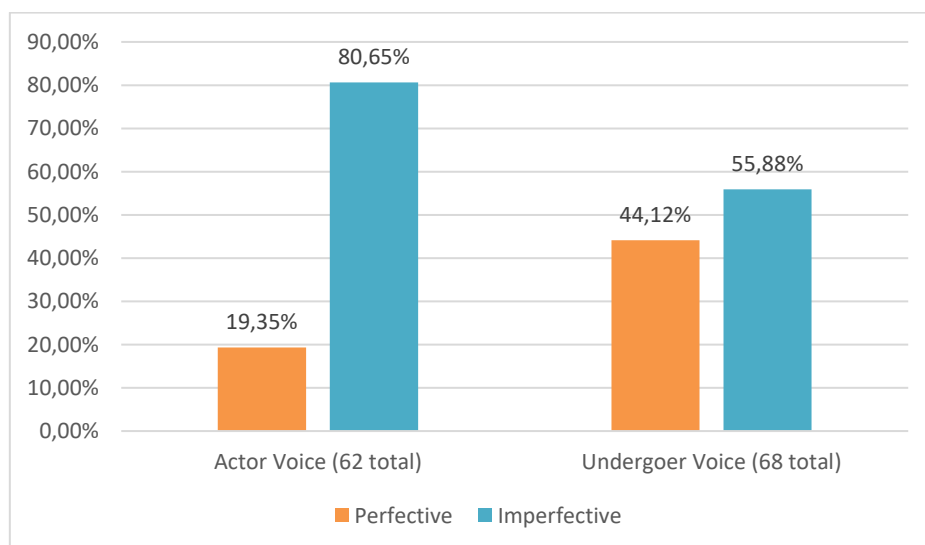


Figure 4.5: Aspect marking and voice

While aspect is fairly evenly distributed in the UV-clauses, imperfective aspect is much more common than perfective in AV-clauses. Seen from another perspective, of the 42 perfective clauses, 12 (28.57%) were AV-clauses while 30 (71.43%) were UV-clauses. Of the 88 imperfective clauses, 50 (56.82%) were AV-clauses while 38 (43.18%) were UV-clauses. These data are summarized in figure 4.6.

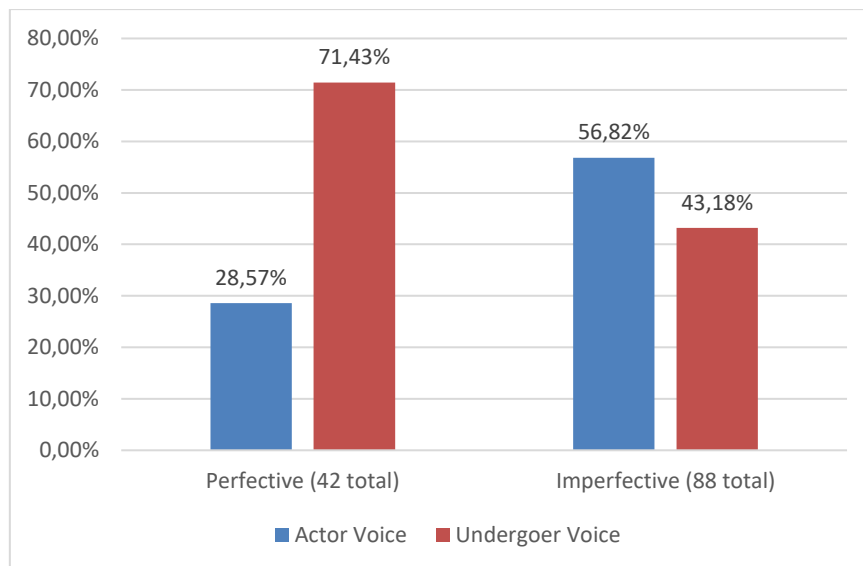


Figure 4.6: Aspect marking and voice

These data seem to indicate that while imperfective aspect does not motivate voice choice in Äiwoo, perfective aspect does, so that, if a clause appears with perfective aspect, undergoer voice is more likely to be selected.

## Punctuality

Concerning punctuality, all 160 clauses were coded for either durative or punctual. Of the 80 AV-clauses, 71 (88.75%) were coded as durative, while 9 (11.25%) were coded as punctual. Of the 80 UV-clauses 70 (87.50%) were coded as durative while 10 (12.50%) were coded as punctual. Figures 4.7 summarizes these data.

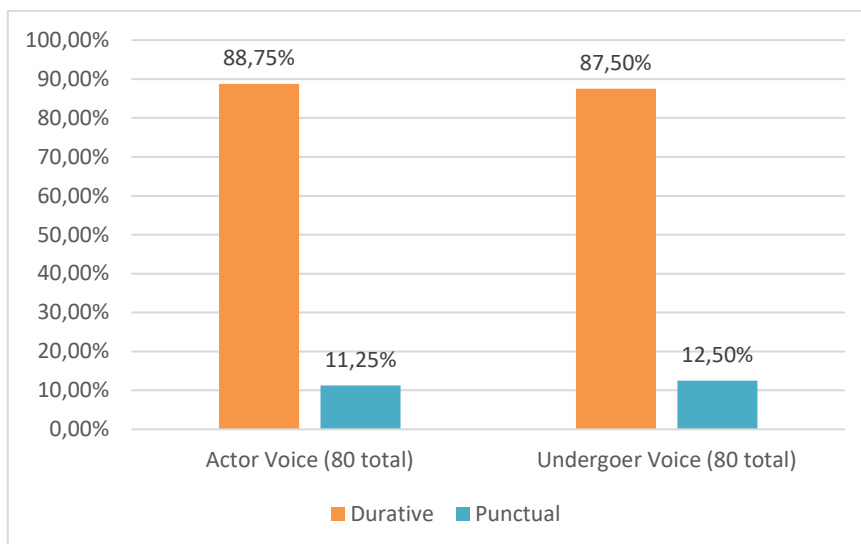


Figure 4.7: Punctuality and voice

These data indicate that there is no difference in punctuality between the AV- and UV-clauses. Consequently, this parameter does not seem to be a motivating factor for voice choice in Äiwoo, based on these data.

### Volitionality

All 160 clauses were coded as either *volitional* or *non-volitional*. 78 (97.50%) of the 80 AV-clauses were coded as volitional while only 2 (2.50%) were coded as non-volitional. The pattern was identical for the UV-clauses. In all of the four non-volitional clauses, the predicate was ‘give birth’. In section 3.2.2, I discussed the difficulty of deciding whether this predicate should be coded as volitional or non-volitional but argued that because it is not an action that the actor can choose to set in motion at her own volition, it should be coded as non-volitional. It is of little consequence, however, whether these four clauses are coded as volitional or non-volitional. The data clearly show that volitional actions are by far the most common, regardless of voice construction. These data are summarized in figure 4.8:

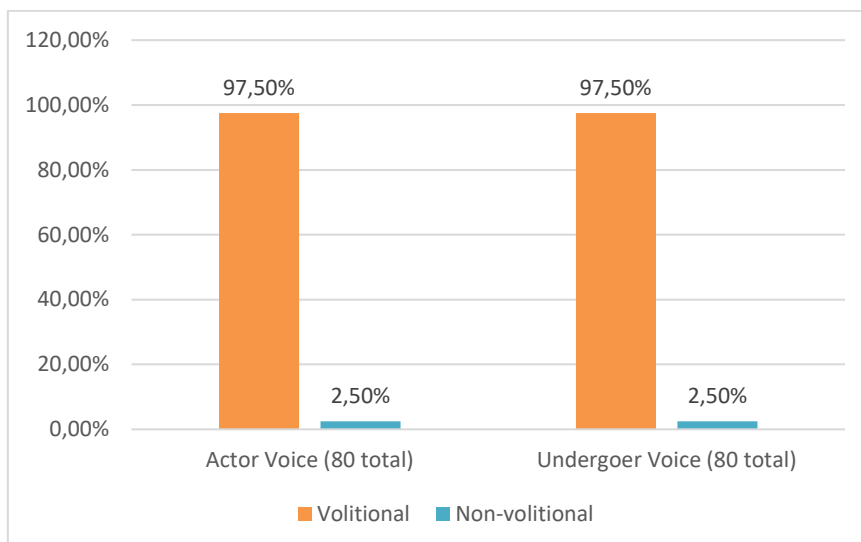


Figure 4.8: Volitionality and voice

As with punctuality, there is no difference in volitionality between the AV- and UV-clauses. Consequently, there are not enough non-volitional clauses in my dataset to conclude whether or not this parameter is a motivating factor for voice in Äiwoo.

### Affectedness of O

Concerning affectedness of O, all undergoer arguments were coded as either *affected*, *partially affected* or *not affected*. However, as there were only 4 instances of non-affected undergoers (3 in UV and 1 in AV), this aspect was omitted from my analysis, as it was too infrequent to indicate anything about voice choice. Concerning the remaining two, 48 (60.76%) of the remaining 79 AV-clauses had fully affected undergoers, while 31 (39.24%) had partially affected undergoers.

Of the remaining 77 UV-clauses, 40 (51.95%) had fully affected undergoers, while 37 (48.05%) had partially affected undergoers. Figure 4.9 summarizes these results.

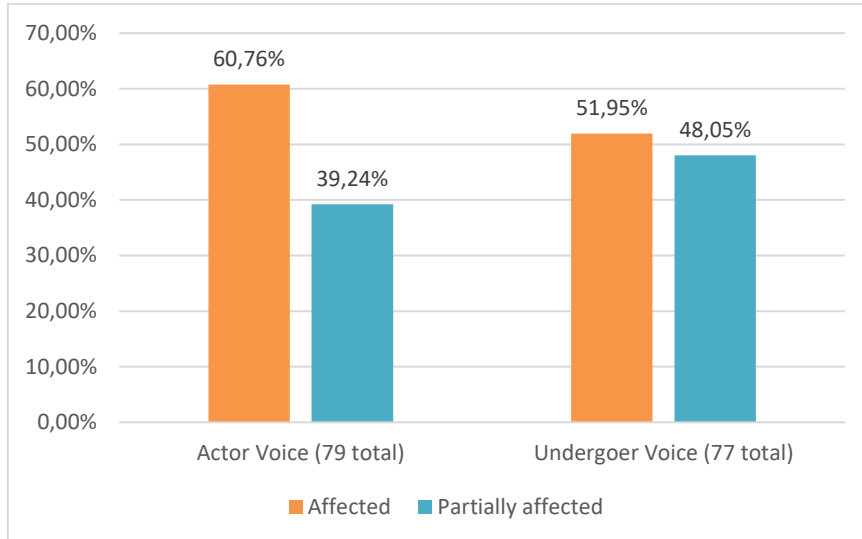


Figure 4.9: Affectedness of O and voice

Seen from the other perspective, of the 88 fully affected undergoers, 48 (54.55%) were found in AV-clauses while 40 (45.45%) were found in UV-clauses. Of the 68 partially affected undergoers, 31 (45.59%) were found in AV-clauses while 37 (54.41%) were found in UV-clauses. These data are presented in figure 4.10.

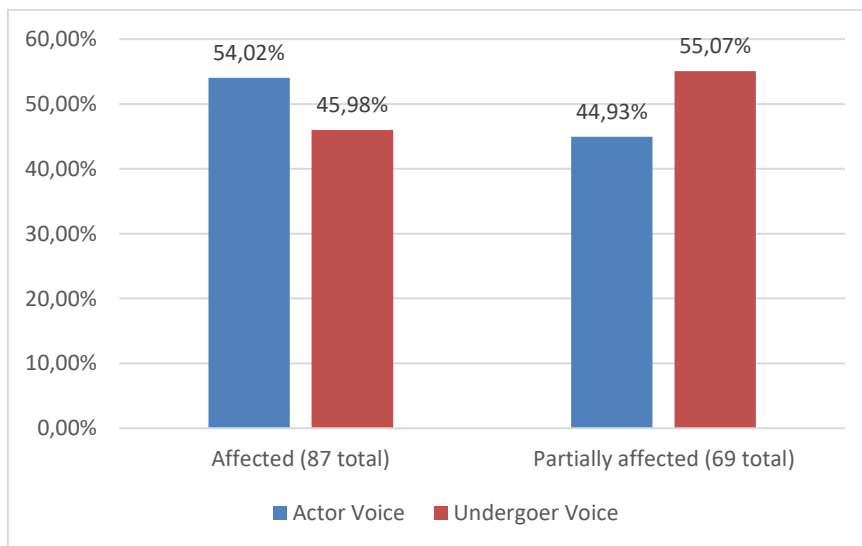


Figure 4.10: Affectedness of O and voice

The data show a mirror image between affectedness in the AV- and UV-clauses where the former had slightly more fully affected undergoers while the latter had slightly more partially



affected undergoers. While there is some pattern between affectedness of O and voice, I show in section 4.2 that this pattern is not statistically significant.

### Individuation of O

Next, undergoer arguments in each of the 160 clauses were coded for individuation with the values *general* or *particular*. In the 80 AV-clauses, 29 (36.25%) of the undergoers were coded as general while 51 (63.75%) were coded as particular. In the 80 UV-clauses, 17 (21.25%) of the undergoers were coded as general, while 63 (78.25%) were coded as particular. These results are illustrated in figure 4.11.

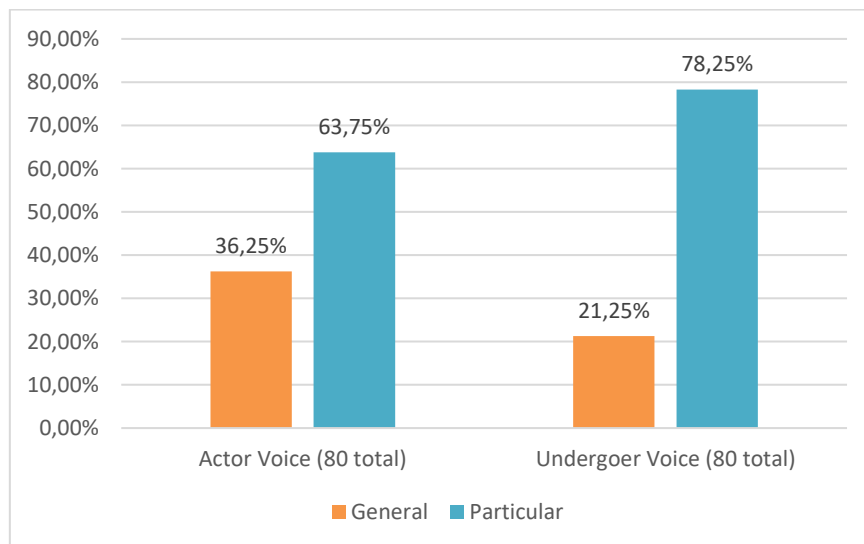


Figure 4.11: Generality of O and voice

From figure 4.11, it is evident that particular undergoers are more common than general ones in clauses of either voice. However, the difference between general and particular is even greater in undergoer voice than in actor voice.

From another perspective, 46 (28.75%) of the total 160 undergoers were coded as general, while 114 (71.25%) were coded as particular. Of the 46 coded as general, 29 (63.04%) occurred in AV-clauses while 17 (36.96%) occurred in UV-clauses. On the other hand, 51 (44.74%) of the 114 undergoers coded as particular occurred in AV-clauses while 63 (55.26%) occurred in UV-clauses. These data are summarized in figure 4.12.

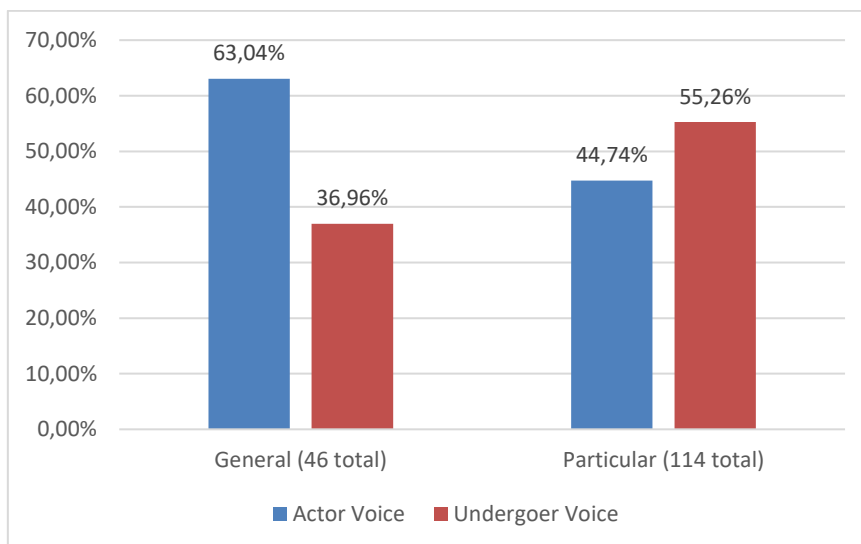


Figure 4.12: Generality of O and voice

It becomes evident from figure 4.12 that general undergoers in AV-clauses are nearly twice as frequent as general undergoers in UV-clauses. The difference is not quite as striking concerning particular undergoers. Nevertheless, they occur more frequently in UV-clauses. Like the previous parameter, affectedness of O, the apparent pattern illustrated by these data is found not to be statistically significant, as discussed in section 4.2.

## Animacy

The final transitivity parameter I investigated was animacy. I coded the actors and undergoers of all 160 clauses for either animate or inanimate and found only one clause with an inanimate actor argument, which is illustrated in (60).

- (60) *ngaa lâto nyinââ ku-wo-mă tevăivă lâ*  
 so then wave IPFV-go-DIR1 stone DEIC.DIST  
*ku-tu=nâ*  
 IPFV-bring=DEIC.DIST  
 ‘So the waves come and scatter the stones.’ (Fishing cairn 367)

Because there were only one instance of an inanimate actor, there was insufficient evidence to conclude whether or not the animacy of actors is a motivating factor for voice choice in Äiwoo. Concerning the animacy of undergoers, the results showed that inanimate undergoers were far more frequent than animate undergoers. Of all 160 clauses, 31 (19.38%) had animate

undergoers while 129 (80.63%) had inanimate undergoers.

Of the 31 clauses with animate undergoers, 14 (45.16%) were found in AV-clauses while 17 (54.84%) were found in UV-clauses. Of the 129 clauses with inanimate undergoers, 66 (51.16%) were found in AV-clauses while 63 (48.84%) were found in UV-clauses. Seen from another perspective, 14 (17.50%) of the 80 AV-clauses had animate undergoers, while 66 (82.50%) had inanimate undergoers. Of the 80 UV-clauses, 17 (21.25%) had animate undergoers, while 63 (78.75%) had inanimate undergoers. These results are summarized in figure 4.13.

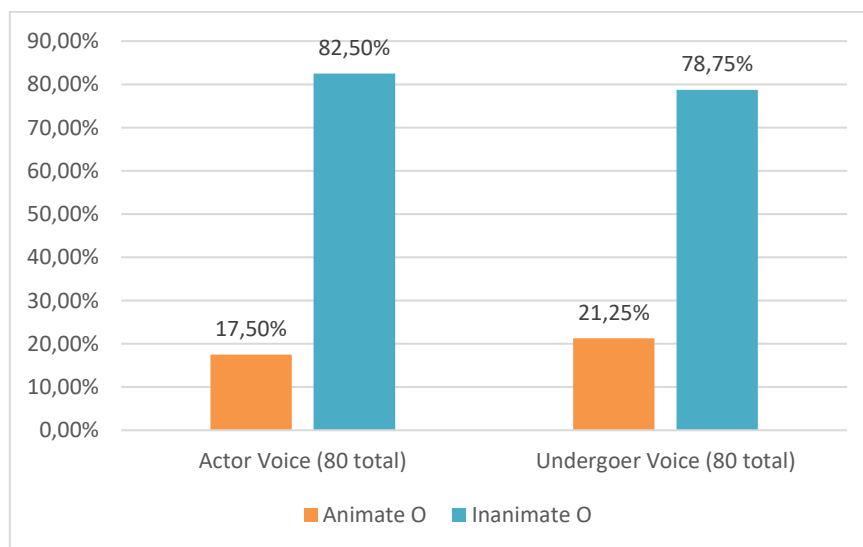


Figure 4.13: Animacy of O and voice

Based on these data, there is no reason to assume that animacy of the undergoer motivates voice choice in Äiwoo. The logistic regression test presented in the next section confirms this.

Actors and undergoers of all 160 clauses were also coded for either SAP (speech act participant, i.e. first- or second person referent) or third person referent. None of the undergoer arguments in either voice construction was a SAP. Consequently, the person marking of undergoers in these data provides no information about voice choice in Äiwoo. Concerning actors then, 42 (52.50%) of the actors in the 80 AV-clauses were SAPs while 38 (47.50%) were third person referents. The pattern was almost identical in undergoer voice: 44 (55%) of the actors in the 80 UV-clauses were SAPs, while 36 (45%) were third person referents. These data are summarized in figure 4.14.

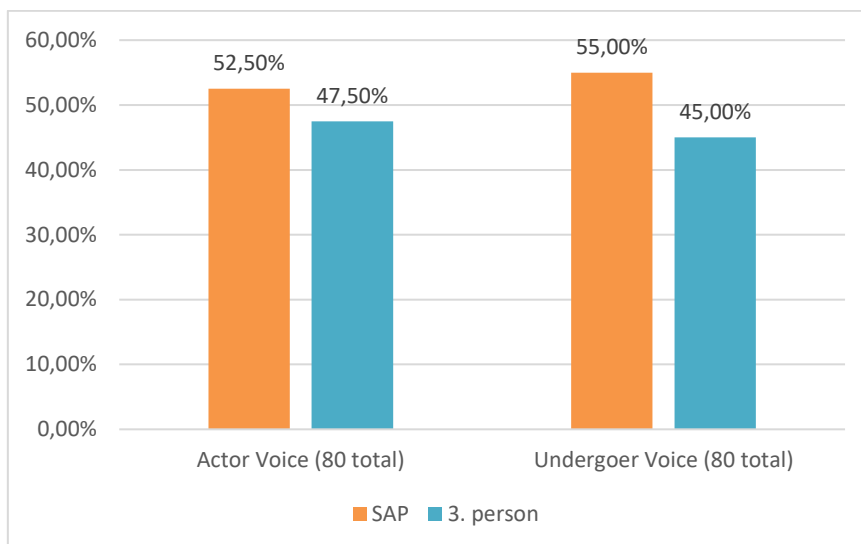


Figure 4.14: SAP and voice

As with animacy, there is no difference in SAP vs. third person referents in the AV- and UV- clauses. Consequently, whether the actor is a SAP or not does not seem to influence voice choice in Äiwoo, based on these data. This is confirmed by the results from the logistic regression test discussed in section 4.2.

### 4.1.3 Activation states

The third framework I applied was the three different activation states as defined by Chafe (1987). I coded all of the 160 clauses for activation state for both actor and undergoer arguments. I present the results in this section. First, actor arguments of either voice construction were almost without exception coded as active. Of the 80 AV-clauses, only two of the actor arguments were not active. Both of these were coded as semi-active. The first of these is given in (61).

- (61) Lâ tumwä i-lâwââ-kä=to opo nyigi lakwaio  
 DEIC:DIST father.3MIN PFV-build-DIR3=now house one small  
 tä=nâ  
 3MIN.POSS.LOC=DEIC:DIST  
 ‘Then her father built a small house for her.’ (Moliki 78)

The actor argument in this clause is Moliki’s father. This referent has not been mentioned within the eight preceding clauses and the focus of the discourse has been Moliki herself,

which has been active in the meantime. However, throughout the narrative the listener has become well acquainted with the fact that Moliki has a father who is a central part of her life. In other words, Moliki’s father is accessible in the mind of the listener. Because of this, this referent is coded as semi-active.

The other instance of a semi-active actor is shown in (62).

- (62) eâmo gelitivo-ngopu läto ki-li-eange-mä nâbo=kâ  
 then mother.1MIN-1AUG then IPFV-3AUG-slice-DIR1 dried.breadfruit=DEIC.DIST  
 ‘then our mothers start cutting chips for nâbo (dried breadfruit).’  
 (Fermenting breadfruit 5, 6)

The actor argument is *our mothers*. While this is the very first mention of this referent, it can safely be assumed that the both the speaker and addressee know who their mothers are. As such, this noun phrase contains accessible information and is coded as semi-active.

Of the 80 UV-clauses, all of the actor arguments were coded as active. Because actor arguments are almost exclusively active, regardless of voice, the activation state of actors does not seem to be a motivating factor for voice choice, at least according to my dataset. Concerning undergoers, however, the results seem to show a clear correlation between activation state and voice. In the AV-clauses, 34 (42.50%) of the undergoer arguments were coded as active, 13 (16.25%) as semi-active and 33 (41.25%) as inactive. In the UV-clauses, 66 (82.50%) of the undergoer arguments were coded as active, 9 (11.25%) as semi-active and only 5 (6.25%) as inactive. This pattern is illustrated in figure 4.15.

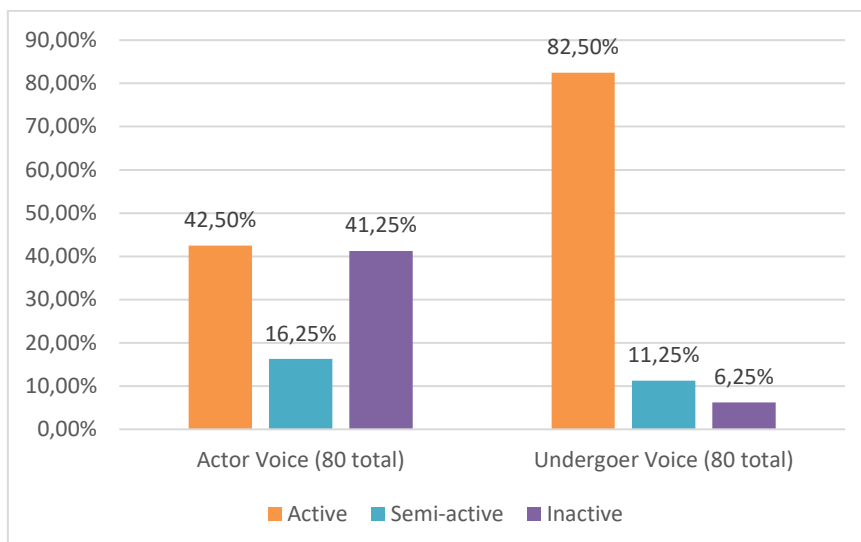


Figure 4.15: Activation state of O and voice

Seen from another perspective, of all the 100 undergoer arguments coded as active, 34 (34%) occurred in AV-clauses while 66 (66%) were found in undergoer voice. Of all the 22 undergoer arguments coded as semi-active, 13 (59.09%) were found in actor voice construction while 9 (40.91%) occurred in undergoer voice. Finally, of all the 38 undergoer arguments coded as inactive, 33 (86.84%) were found in AV-clauses while only 5 (13.16%) were found in undergoer voice. This pattern is illustrated in figure 4.16.

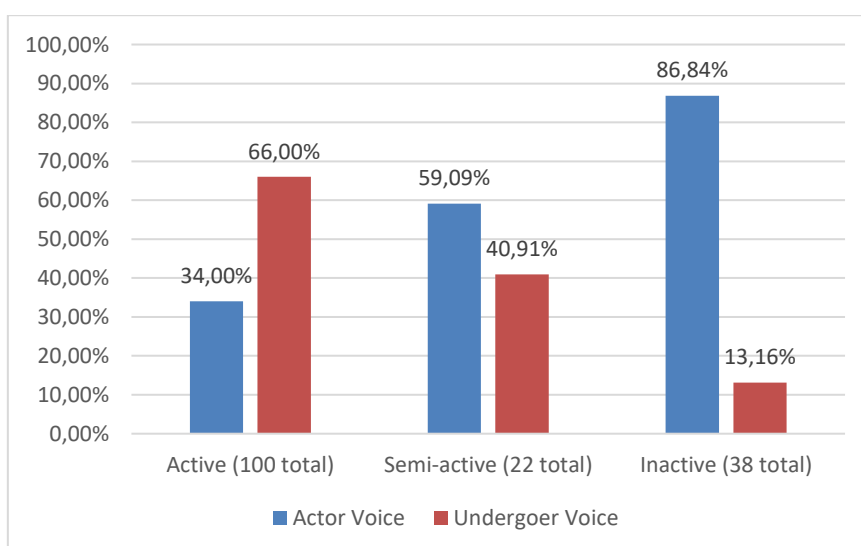


Figure 4.16: Activation state of O and voice

Two observations are striking from these results. First, the great majority of undergoer arguments found in UV-clauses are active, as seen in figure 4.15. Second, the majority of inactive undergoer arguments occur in AV-clauses, as seen in figure 4.16.

This pattern suggests that when the undergoer is an inactive concept, actor voice is more likely to be selected. When the undergoer is an active concept however, undergoer voice is more likely to be selected. This can be illustrated in the following examples:

- (63)    *iveveto*                    *singeda*                    *näsäpelivanoi*                    *jii*  
 PFV-buy=now                    woman                    IRR-wife-3AUG                    3AUG  
 ‘he bought wives for them’
- Ive*                    *singeda*                    *eve,*                    *näsäpelivanogui*                    *ilâ*  
 PFV-pay                    woman                    three                    IRR-wife-3MIN-3AUG                    that  
*pelivalipiänâ*  
 brother.in.law.3MIN=DEIC.DIST  
 ‘He bought three women, for his brothers-in-law to marry.’
- Ä                    *iveepu*                    *nyigi epu*                    *singeda*                    *näsäpelivano*                    *gino*  
 and                    PFV-pay.also                    one                    also                    woman                    IRR-wife                    son  
 ‘And he bought another one to be his son’s wife.’                    (Settlement 145-47)

In the first clause in (63), the argument *singeda* ‘woman’ is introduced and is therefore inactive. In this clause, actor voice is selected, signalled by the AV-form of the verb *veve*. *Singeda* ‘woman’ is repeated in each of the subsequent clauses and is now active. In both cases, undergoer voice is selected, signalled by the UV-form of the verb *ve*. This clearly illustrates that UV-clauses are used when the undergoer represents active concepts. Because UV-clauses are by far the most common transitive clauses, one might assume that this is the unmarked construction to use when referring to an undergoer that is active concept. When introducing new information in undergoer position, either intransitives or AV-clauses are used. However, because 38.46% of all undergoer arguments in actor voice do not represent inactive concepts, being coded as either active or semi-active, signalling inactive concepts cannot be the only function of the AV-construction.

## 4.2 Significance testing

In order to see if the correlations found in section 4.1 are significant, I performed a logistic regression test. The response variable is *voice* with the values *AV* and *UV*. The predictor variables are *aspect*, *affectedness of O*, *individuation of O*, *activation state of O*, *referential distance of A*, *topic persistence of A*, *referential distance of O* and *topic persistence of O*.

The preliminary test showed that the only predictor variables with a significant correlation to voice are aspect and activation state of O. Perfective aspect is significant with a p-value of 0.008, meaning this variable is significant with an alpha-level of .01. The effect size shows that undergoer voice is 3.8 times more likely than actor voice to be chosen if aspect is perfective rather than imperfective. Furthermore, the results showed that the inactive value of the variable activation state of O is significant with a p-value of 0.0003, meaning it is significant with an alpha level of .001. The effect size for this variable shows that actor voice is 11.92 times more likely than undergoer voice to be chosen if the activation state of O is inactive rather than active. In short, perfective aspect makes undergoer voice almost four times as likely to be selected, while inactive undergoers make actor voice almost 12 times as likely to be selected according to the results of the logistic regression test.

However, because 8 different variables were tested simultaneously, there is a chance that the significant results for aspect and activation state of O are affected by interactions with the other variables. In order to see if these variables were still significant without being affected by other variables, I performed the same test with only aspect and activation state of O as predictor variables. Furthermore, I performed an interaction test between these two predictor variables in order to see if they affect each other. The results showed 1) that there was no significant interaction between aspect and activation state and 2) that both aspect and activation state are still statistically significant without being affected by interaction from other variables. The p-value for perfective aspect is 0.004, meaning it is still significant with an alpha-level of .05. Likewise, inactive activation state of O is still significant with an alpha-level of .001 with a p-value of 1.31e-06. Furthermore, the effect sizes show that undergoer voice is still 3.8 times more likely than actor voice to be selected if aspect is perfective rather than imperfective and that actor voice is now 15.01 times more likely than undergoer voice to be selected if the activation state of O is inactive rather than active. In sum, both aspect and affectedness of O are significant predictors of the response variable voice.



### 4.3 Discussion on statistical analysis

I have chosen to perform statistical significance tests in order to see if the tendencies discovered in my sample can be said to exist in the population as well. However, there is a potential problem with applying significance tests to my dataset. This problem concerns independence of observations, which is a requirement that must be met when performing all tests of statistical significance. Jensen (2018: 461) argues that for the observations in a dataset to be independent of each other, no two observations can come from the same speaker. In other words, each speaker in a dataset must be the source of one and only one observation. This seems to be a common view by statisticians and researchers who use statistics in linguistics. Rietveld et al. (2004) argue that the use of the same speaker or writer may give rise to so-called unit dependencies where *unit* refers to the speakers, writers or texts. They write,

Disregarding the actual units of sampling in processing frequencies, however, may cause serious problems, since the observations from the same speaker, writer, or text may be more similar than those from different ones. Disregard of the unit or level of sampling can only imply the violation of the assumption of independence (Rietveld et al. (2004: 352).

In other words, if a unit (speaker, writer or text) has their own unique way of expressing themselves, this difference will not be accounted for if you take more than one observation from the same unit. Consequently, Howell (2007: 152-53) argues that the best way to avoid violating the assumption of independence of observation is “to make certain that the total of all observations (N) equals precisely the number of participants in the experiment.”

Based on these arguments, it is possible to argue that not all 160 clauses in my dataset can be considered independent of each other, because a great number of them come from the same speakers. There are 23 different texts in my dataset. Two of these contain speech from the same speaker. This means that I only have 22 different speakers and consequently 22 different independent observations. It can therefore be argued that, if voice in Äiwoo were a highly varying phenomenon, the use of all 160 observations in significance testing could lead to wrong conclusions concerning voice choice.

However, it is reasonable to assume that this view can be rejected based on the following arguments. Because Jensen’s (2018) arguments concern second language research, two observations from different speakers are expected to vary for several reasons, like different linguistic backgrounds, number of second languages previously learned, amount of time they have spent learning the target language, et cetera. Because of these factors, each

speaker will have their unique way of speaking in their target language and consequently no two observations from the same speaker should be treated as independent of each other. On the other hand, the speakers in my dataset are all native speakers of the same language: Äiwoo. It can reasonably be assumed that voice, along with aspect and other morphosyntactic properties, is an inherent part of the linguistic structure of Äiwoo. As such, there is no reason to assume that voice varies greatly between the speakers in my dataset. Consequently, the 160 observations in my dataset can be treated as independent after all. A quote from Rietveld & van Hout (2005: 127) seems to support this view. They write:

On many occasions it is defensible that a researcher treats violations of statistical assumptions as a harmless nuisance. One should, however, always ask oneself what kind of properties in the data studied can cause violations of underlying assumptions and try to answer the question whether the violations found are inherent to the phenomenon under investigation. One should query whether these properties can be accounted for in the model which underlies the data. This implies that one should always inspect the data thoroughly for possible trouble sources, however negligible they may seem in testing specific hypotheses.

In short, certain violations of assumptions can be justified, so long as these violations do not alter the results of the tests. Again, there is no reason to believe that voice choice in Äiwoo is a linguistic phenomenon that varies greatly between speakers. On the contrary, it is a phenomenon entrenched in the linguistic structure of the language.

Finally, I am not alone in applying significance tests to datasets of this kind. As discussed in chapter 2, Wouk (1996) investigated correlations between voice and other linguistic factors in Spoken Jakartan Indonesian (SJI). Her dataset consisted of 88 clauses of each voice construction (actor voice and undergoer voice), which came from tape-recorded conversations of 15 speakers. Despite the fact that the source of her 176 clauses was 15 speakers, i.e. that the number of observations did not equal the number of units, she still applied chi-square tests on her dataset and found significant correlations between voice and mood, voice and punctuality and voice and referentiality. While the fact that another researcher has made this choice is not in itself evidence of its legitimacy, it does support my decision to reject the arguments against it. In sum, based on the arguments above, the assumption that the 160 sentences in my dataset can be treated as independent observations is maintained and consequently that the significance tests performed are valid.

In conclusion, of all the variables investigated, only aspect and activation state of O were shown to be significant predictors of voice in Äiwoo. Concerning aspect, the results show that undergoer voice is almost four times (3.8) more likely than actor voice to be chosen

if the aspect is perfective rather than imperfective. Concerning activation state, the results show that actor voice is approximately 15 times (15.01) more likely to be chosen if the activation state of the undergoer is inactive rather than active. In the final chapter of this thesis (chapter 5), I will discuss what consequences these results have for voice choice in Äiwoo.

## 5 Summary and discussion

Äiwoo is a language with a so-called symmetrical voice system, where two or more transitive constructions alternate according to which semantic argument receives the status as the most prominent syntactic argument. This type of system has been reconstructed for Proto Austronesian and Proto Malayo-Polynesian and is found in a large number of western Austronesian languages. However, the system has not been reconstructed for Proto Oceanic and most Oceanic languages lack voice systems. Instead, clauses in these languages alternate according to a difference in transitivity. Äiwoo, being an Oceanic language, differs from other Oceanic languages in having a symmetrical voice system, rather than a system of transitivity alternations. The aim of this thesis is to investigate which factors motivate a speaker of Äiwoo to choose one of the two main voice constructions, actor voice or undergoer voice, over the other. My dataset consists of 160 clauses spoken by native speakers of Äiwoo. In order to see what factors motivate voice choice in Äiwoo, I adopted three different methodological frameworks: topic continuity developed by Givón (1983, 1994), discourse transitivity developed by Hopper & Thompson (1980), and the three activation states as discussed by Chafe (1987). I will now summarize and discuss the results found for each of the three frameworks.

### 5.1 Summary

#### 5.1.1 Topic continuity

Concerning topic continuity, I coded the actor and undergoer arguments of all 160 clauses for the measurements referential distance (RD) and topic persistence (TP), both of which had a maximum upper limit of 10 units. While a low value for referential distance corresponds to a high degree of continuity, for topic persistence, a high value corresponds to a high degree of continuity. The motivation for this coding process was to find out whether voice choice is governed by the topicality or continuity of the voice-selected argument and to test Næss' (2015) hypothesis that the topic persistence of the voice-selected argument motivates the speakers' choice of voice.

The results showed that there was no difference in the continuity of the actor argument according to either measurement. Furthermore, the results for topic persistence of both actor

and undergoer arguments as well as the results for referential distance of actor arguments were the same for both actor voice and undergoer voice. Consequently, Næss' (2015) hypothesis that voice is governed by the thematic importance, i.e. topic persistence of the voice-selected argument can be rejected. The only value that differed between actor voice and undergoer voice was the referential distance of the undergoer arguments (RD O). In actor voice, the median value for RD O was 4 clauses while in undergoer voice, the median value was 1 clauses. Because a low value for referential distance corresponds to a higher degree of continuity, this suggests that the continuity of the undergoer argument is higher in undergoer voice than in actor voice. What does this imply for voice choice? As shown in the previous chapter, the results from the logistic regression test showed no significant correlation between voice and referential distance of the undergoer argument. Because the correlation was not significant, I argue in section 5.1.4 that this apparent pattern is merely a coincidence of the results of another correlation, namely between voice and activation state of O.

### **5.1.2 Discourse transitivity**

Next, using to the framework by Hopper & Thompson, I coded all 160 clauses for the following parameters: kinesis, aspect, punctuality, volitionality, affectedness of O, individuation of O and animacy. All of these parameters correspond to a high degree of semantic transitivity, according to Hopper & Thompson (1980). Thus, if correlations between voice and several of these parameters were to be found, one could argue that the voice choice in Äiwoo is governed by degree of semantic transitivity, so that one voice construction has a higher degree of semantic transitivity than the other.

The results showed that the following parameters were the same regardless of voice: kinesis, punctuality, volitionality and animacy. Concerning two of the remaining three, affectedness of O and individuation of O, the data suggested that there was some correlation between these parameters and voice. The results from the logistic regression test however, proved these correlations not to be statistically significant. That leaves only one parameter, namely aspect. From the results presented in the previous chapter, two observations can be made concerning this parameter and voice. First, 80.65% of all aspectually marked AV-clauses were imperfective. Of the aspectually marked UV-clauses, by comparison, only 55.88% were imperfective, suggesting that aspect is more evenly distributed in undergoer voice. Second, of the 42 perfective clauses, 71.43% were UV-clauses, while only 28.57% were AV-clauses. Of the 88 imperfective clauses, on the other hand, 56.82% were AV-

clauses, while 43.18% were UV-clauses. These observations suggest there is in fact a correlation between aspect and voice in Äiwoo. Indeed, the logistic regression test presented in chapter 4 showed this to be the case. I repeat the results here: If the aspect of a clause is perfective, undergoer voice is almost four times as likely to be selected. In Hopper & Thompson's transitivity framework, imperfective aspect corresponds to a lower degree of semantic transitivity while perfective aspect corresponds to a higher degree of semantic transitivity. This means that according to this parameter alone, undergoer voice is more semantically transitive than actor voice.

In chapter 1, section 1.5, I explored the possibilities of analysing AV-clauses as transitive clauses with discord, as defined by Margetts (2008). The reasoning behind this was that AV-clauses mark their actors with a prefix on the verb just like intransitive clauses do. Thus, like transitive discord clauses in Oceanic languages like Manam and Saliba, AV-clauses in Äiwoo appear to be transitive from a syntactic point of view, because of the presence of two arguments, but intransitive from a morphological point of view, because of their intransitive head marking. UV-clauses however, mark their actors with a suffix, rather than a prefix, on the verb and hence differ from both intransitive clauses and AV-clauses. The alternative analysis of the Äiwoo clause system presented in (32) is repeated here in (64):

- (64) One-argument clause with actor **prefix** = intransitive clause  
 Two-argument clause with actor **prefix** = transitive clause with discord  
 Two-argument clause with actor **suffix** = transitive clause with accord

This analysis states that two-argument clauses with actor prefix, i.e. what has been referred to as AV-clauses are actually transitive clauses with discord, while two argument clauses with actor suffix, i.e. what has been referred to as UV-clauses are actually transitive clauses with accord. However, as already discussed in chapter 1, this only makes sense from a morphosyntactic perspective. Furthermore, if we assume that two morphosyntactically different constructions in a given language can be explained in terms of transitivity, we would expect these constructions to also differ with regard to semantic transitivity, as defined by Hopper & Thompson. Hence, if AV- and UV-clauses were to be analysed as discord and accord clauses respectively, they would be expected to differ in terms of their semantic transitivity, in particular with regard to the individuation of the undergoer. As became clear in the previous chapter however, while undergoers were slightly more individuated in UV-

clauses than in AV-clauses, this difference was not statistically significant. Furthermore, while this parameter also proved not to be significant, undergoers were, interestingly, found to be more affected in AV-clauses than in UV-clauses. This observation speaks against the claim that AV-clauses are less semantically transitive than UV-clauses. Finally, undergoers were found to be equally inanimate in AV- and UV-clauses. Hence, none of the parameters that deals with the status of the undergoer was found to be significantly different in the two constructions.<sup>24</sup> Likewise, none of the parameters that deal with the status of the clause proved to be significantly different between the two clauses either, with the exception of aspect. Recall that undergoer voice is 3.8 times more likely than actor voice to be selected if aspect is perfective rather than imperfective. Because perfective aspect corresponds to a higher degree of semantic transitivity according to Hopper & Thompson (1980), it can safely be argued that UV-clauses are more transitive than AV-clauses from a semantic perspective. However, Margetts (2008: 42) clearly states that it is the properties of the object/undergoer that determines the choice between a canonical transitive clause and a discord clause in Oceanic languages. Aspect is a property of the clause as a whole and not of the undergoer in particular. Hence, the correlation between UV-clauses and perfective aspect is not sufficient evidence to analyse AV-clauses as discord clauses.

### **5.1.3 Activation state**

Finally, the undergoer arguments of all 160 clauses were coded for activation state with three values: active, semi-active and inactive. I repeat the results from the data in chapter 3 here. Of all the AV-clauses, 42.50% had active undergoers, 16.25% had semi-active undergoers, while 41.25% had inactive undergoers. Of the UV-clauses, 82.50% had active undergoers, 11.25% had semi-active undergoers, while only 6.25% had inactive undergoers. From another perspective, of the 100 active undergoers, 34% occurred in AV-clauses, while 66% occurred in UV-clauses. Of the 22 semi-active undergoers, 59.09% occurred in AV-clauses, while 40.91 occurred in UV-clauses. Finally, of the 38 inactive undergoers, 86.84% occurred in AV-clauses, while only 13.16 occurred in UV-clauses. From these data, two main observations arise. First, there is a tendency for inactive undergoers to occur in AV-clauses. Second, there is a tendency for UV-clauses to host active undergoers. The results from the logistic regression test showed that there was indeed a significant correlation between

---

<sup>24</sup> Though in section 5.3, I argue that the activation state of undergoers determines their degree of individuation.

activation state and voice. As stated in section 3.4, actor voice is 15 times more likely than undergoer voice to be selected if the undergoer is inactive than if it is active.

Based on these observations, I propose that one of the functions of actor voice is to introduce new undergoer arguments into the discourse. Once these arguments have been introduced and then are given, the speaker chooses to use undergoer voice until a new undergoer argument is to be introduced. The question then arises: Where are actor arguments introduced? In my dataset, there were no inactive actor arguments and consequently no introduction of new actors into the discourse. This observation is in line with what Du Bois (1987) terms Preferred Argument Structure or PAS. According to PAS, only one new referent is introduced in a single clause and that referent is typically introduced in either S or O position (Du Bois 1987: 828). This means that new referents are either the single argument of an intransitive clause, or the undergoer argument of a two-argument clause. Consequently, Du Bois' notion of PAS explains why there are no new actors in my dataset: it does not contain any intransitive clauses and consequently no S position for actors to be introduced. It does not, however, explain why new undergoers are significantly more frequent in actor voice than in undergoer voice. Again, my hypothesis here is that one of the functions of actor voice is to introduce new referents into the discourse.

What is interesting is that Du Bois' data come from the Mayan language Sacapultec Maya, which has ergative morphology and accusative syntax. Hence, the morphosyntax of this language differs greatly from the symmetrical voice system found in Äiwoo. To the best of my knowledge, the investigation into PAS has never been done on any languages with a symmetrical voice system. Hence, we know very little about how symmetrical voice interacts with preferred argument structure as defined by Du Bois (1987). In Sacapultec Maya, like in most languages that do not have a symmetrical voice system, a new argument can occur in either S position of intransitive clauses or in A or O position of transitive clauses. Hence, there are three possible positions across two clause types for a new argument to occur. In Äiwoo however, a new argument can occur in S position of intransitive clauses, in A or O position of AV-clauses, or in A or O position of UV-clauses. Hence, there are five different possible positions across three different clause types for a new argument to occur in Äiwoo, as summarized in (65).



- (65) Intransitive clause: S  
 AV-clause: A O  
 UV-clause: A O

The results from the present study indicate that in a preferred argument structure in Äiwoo, the O position of AV-clauses is one of the positions where new arguments typically occur. I have not investigated intransitive clauses in this study, but if the PAS of Äiwoo is anything like that in Sacapultec Maya, then the S position of these clauses might also be a possible position for new arguments to occur.

In short, because languages with a symmetrical voice system have three or more clause types, it would be interesting to conduct further research into the preferred argument structure of these languages. In particular, further research might be able to investigate whether new arguments are typically introduced in S position of intransitive clauses and in O position of AV-clauses.

#### **5.1.4 Topic continuity vs. activation state**

Three frameworks have been discussed in this thesis: topic continuity, discourse transitivity and activation states. Concerning topic continuity, the results showed some notable differences between undergoer voice and actor voice in the RD of undergoers. The median value for RD was 4 in actor voice and 1 in undergoer voice. However, while this difference is notable, it is not statistically significant, as shown by the logistic regression test. Concerning the activation states of O, the results show that there was indeed a significant correlation between this parameter and voice, such that actor voice is 15 times more likely than undergoer voice to be selected if the undergoer is inactive than if it is active.

These parameters are clearly related. Givón (1994: 9) states that the referential distance of an argument determines how cognitively accessible it is. An active concept, on the other hand, is defined by Chafe (1987: 25) as “one that is currently lit up, a concept in a person’s focus of consciousness,” or simply as information that is given. As such, it is difficult to distinguish these parameters from each other and it would be tempting to group them together as one single parameter. However, they do in fact differ in terms of how they are measured. While referential distance or cognitive accessibility is a matter of degree, givenness relies on a binary value: given or new. According to the first parameter, one argument can be more accessible than the other, depending on the value for referential

distance. Concerning givenness however, if both arguments are given, it does not make sense to speak of one argument being ‘more given’ than the other as an argument is either given or not. Hence, because the former is a numerical variable while the other is a categorical one, they will of course yield different results. Still, the two parameters are clearly related as is evident from the similarities in the results. I therefore suggest that the notable, but not significant, correlation between referential distance of undergoers and undergoer voice is a consequence of the significant correlation between givenness and voice. If undergoers are new, it follows that there has been a considerable amount of time since their last mention. Hence, their referential distance would be high. If, however, undergoers are given, it follows that their previous mention is fairly recent. Hence, their referential distance would be low. This also explains why the referential distance of actors in both AV and UV is low. As already mentioned, actor arguments were almost exclusively coded as active, or given, regardless of voice. Since they are given, then one would of course expect their previous mention to be fairly recent. Hence, their referential distance would be low, which it in fact is. This explains the connection between these two frameworks.

## 5.2 Voice choice in Äiwoo

I return now to the research question: “What factors motivate speakers of Äiwoo to choose one voice construction over the other?” I have found that there is no simple answer to this question. One thing I have found is that none of the factors investigated are conditional for voice, so that they guarantee the choice of one voice construction. Instead, I have found correlations, so that certain factors predict the likelihood of a voice construction being chosen. In particular, I have shown that both aspect and activation state of O significantly correlate with, and consequently predict, the choice of voice in Äiwoo. Perfective aspect correlates with undergoer voice in such a way that undergoer voice is 3.8 times more likely than actor voice to be chosen if the aspect is perfective rather than imperfective. Likewise, inactive undergoers correlate with actor voice in such a way that actor voice is 15 times more likely to be chosen if the undergoer is inactive, rather than active. Hence, the correlation between inactive undergoers and actor voice is even greater than the correlation between perfective aspect and undergoer voice. While these factors clearly explain some of the motivations behind voice choice in Äiwoo, they do not explain the entire picture. There are still several unanswered questions. For instance, if one of the functions of actor voice is to introduce new

undergoers into the discourse, as I suggested in 5.1.3, then why are 42% of the undergoers in the 80 AV-clauses active, rather than inactive? The only answer I can provide at this point is that actor voice must have other functions besides introducing new undergoers into the discourse. The question of what these functions may be, however, remains unanswered. Furthermore, while there is a significant correlation between perfective aspect and undergoer voice, 55.88% of the 68 aspectually marked UV-clauses had imperfective, rather than perfective aspect. This observation suggests that perfective aspect is not the only motivation for the choice of undergoer voice over actor voice. Again, the question of which other factors motivate the choice of undergoer voice remains unanswered.

The lack of answers is not due to lack of efforts to find them. After all, I have adopted not only one, but three different explanatory models. Rather, the lack of answers is arguably due to the simple fact that symmetrical voice is a complex phenomenon. To illustrate this, I refer to a study by Riesberg et al. (2020) which investigates the choice between actor voice and undergoer voice in the Sulawesi language Totoli. In order to investigate what motivates voice in this language, Riesberg et al. look at a number of different factors: animacy, humanness, referential distance, topic persistence, activation states, argument realization, tracking use, generalizability, subordination, mood, valency-increasing morphology, causative and applicative morphology, structural priming, interactivity, and text type. Their dataset consists of 27 texts spanning a period of 2 hours and 50 minutes in total, resulting in 3152 clauses. This dataset comes from 53 different speakers and the genres of the texts are varied. In short, their dataset is of considerable size and their investigation into what motivates voice choice is extensive. Despite this, Riesberg et al. found no conditional factors for voice choice in Totoli. While some of the factors investigated were significantly correlated with voice, they conclude that these factors merely show tendencies. This illustrates that voice choice can be extremely difficult to determine, despite extensive investigations like that conducted by Riesberg et al.

### **5.3 Symmetrical voice vs. transitivity**

As discussed in section 1.5 in chapter 1, it is reasonable to argue that the clause system of Äiwoo is in the process of developing into a transitivity system. The arguments presented there were threefold. First, UV-clauses are vastly more frequent than AV-clauses, by a factor of roughly 7 to 1. Second, several high frequency transitive verbs are only attested with UV-

verb roots. Third, several instances of AV-verb forms being used in UV-constructions are attested, even though these verbs have corresponding UV-verb forms. All of these observations suggest that undergoer voice may eventually become the default transitive clause and consequently, that the symmetrical voice system of Äiwoo might be reanalysed as a transitivity system similar to that found in other Oceanic languages.

When discussing transitivity, it is important to distinguish between morphosyntactic and semantic transitivity. The arguments just discussed all concern morphosyntactic transitivity. As has become evident when discussing the framework of Hopper & Thompson (1980), however, there is a correlation between morphosyntactic transitivity and semantic transitivity. Likewise, there seems to be a correlation between symmetrical voice and semantic transitivity. In most of the studies on voice choice discussed in this thesis, there is a correlation between undergoer voice and a high degree of semantic transitivity. In my study, this is true for aspect. While perfective aspect corresponds to a high degree of semantic transitivity, imperfective aspect corresponds to a low degree of semantic transitivity. This means that the clause type that correlates with perfective aspect will correspond to a higher degree of semantic transitivity according to the parameter aspect. As shown in this thesis, there is a statistically significant correlation between undergoer voice and perfective aspect in Äiwoo. Hence, undergoer voice is semantically more transitive than actor voice according to this parameter. However, none of the other parameters showed the same correlation between undergoer voice and a high degree of semantic transitivity. On the other hand, no parameters showed any correlation between actor voice and a high degree of semantic transitivity either. Hence, the results found in my study do not suggest that actor voice is the semantically more transitive of the two voices. The same correlation is found in the languages discussed in chapter 2. Indeed, in these studies, not only one, but several parameters showed correlations between undergoer voice and a high degree of semantic transitivity. They also showed clear correlations between actor voice and a low degree of semantic transitivity. Specifically, Wouk (1996) found that realis mood, punctual actions and referential undergoers correlate with undergoer voice. Because these factors reflect a higher degree of semantic transitivity according to Hopper & Thompson, undergoer voice is semantically more transitive than actor voice is in SJI, according to these results. Likewise, Wouk (1999) found that undergoer voice correlates with realis mood and referential undergoers in Sasak. Hence, undergoer voice reflects a higher degree of semantic transitivity in this language as well. Finally, McDonnell (2016) found that particular and animate undergoers correlate with undergoer voice in

Besemah. These also correlate with a high degree of semantic transitivity because they constitute what Hopper & Thompson (1980) define as individuated undergoers. Hence, undergoer voice in Besemah also reflects a higher degree of semantic transitivity, according to McDonnell's results.

In sum, these results show that there are at least five different properties that correlate with voice in these languages. Two of these properties, mood and punctuality, are clause-related. The other properties however, relate to the status of the participants. What is most interesting about this is that it is the properties of the undergoer argument, not of the actor argument, that seem to motivate the choice of voice in these languages. Furthermore, whenever these properties of the undergoer reflect a high degree of semantic transitivity, undergoer voice is chosen.

Aside from the parameters discussed by Hopper & Thompson (1980), other participant-related properties have been discussed in this thesis, namely referential distance or cognitive accessibility, topic persistence or thematic importance and activation states. Wouk (1996) found that whenever an argument is both more cognitively accessible and thematically important in SJI, the corresponding voice is chosen. In other words, actor voice is chosen if the actor argument is more cognitively accessible or more thematically important than the undergoer argument and vice versa. Quick (2005) found that if the undergoer is more cognitively accessible or thematically important in Pendau, undergoer voice is chosen. Finally, McDonnell (2016) found that if the undergoer is given, undergoer voice is chosen in Besemah. I argue here that the properties cognitive accessibility and givenness also influence the degree of individuation of the undergoer. If an undergoer is more accessible and given, it is difficult to imagine it not also being referential. Furthermore, because referentiality is one of the properties that define an individuated undergoer according to Hopper & Thompson (1980: 253), cognitively accessible and given undergoers are consequently more individuated than inaccessible and new undergoers. Moreover, because individuated undergoers correspond to a higher degree of semantic transitivity, these parameters also suggest that undergoer voice correlates with a high degree of semantic transitivity in SJI, Pendau and Besemah. Similarly, my results for Äiwoo showed a clear correlation between inactive undergoers and actor voice. Following the same argumentation would then suggest that the undergoers in actor voice are less individuated and furthermore that actor voice reflects a lower degree of semantic transitivity in Äiwoo, compared to undergoer voice, with a higher frequency of given undergoers.

In sum, there is a clear correlation between symmetrical voice and semantic transitivity in the symmetrical voice languages discussed in this thesis, including Äiwoo. What about the relationship between formal or morphosyntactic transitivity and semantic transitivity? One link between symmetrical voice and formal transitivity can be made. One of the parameters of Hopper & Thompson, which has not been discussed in this thesis, is the number of participants. Hopper & Thompson accurately observe that at least two participants are necessary in order for a transfer to take place. As such, a clause with two or more participants corresponds to a higher degree of semantic transitivity than a clause with only one participant (Hopper & Thompson 1980: 252). How a transitive clause is formally encoded is of course language-specific. Pretheoretically however, one thing that clearly distinguishes a transitive clause from an intransitive one cross-linguistically is the presence of a salient undergoer that is separated from whatever entity is the cause or instigator of the action denoted by the clause. This reflects Kemmer's (1993) notion of *relative participant distinguishability*, which refers to how two participants or arguments in a two-argument clause are distinguished from each other, as opposed to being distinguished or individuated from the general background (cf. Hopper & Thompson's individuation parameter). Næss (2007: 23) incorporates this notion with individuation into what she calls *distinctness of participants*, which then refers to "the distinctness of each participant from the general background, and their distinctness from each other in terms of their physical and semantic properties." She further argues that the basic syntactic opposition between an intransitive clause and a transitive clause is reflected by this distinctness of participants (Næss 2007: 23).

In other words, a transitive clause predicates the existence of both an actor and an undergoer. As established in this section, undergoer voice in symmetrical voice languages typically has highly individuated undergoers, while actor voice tends to have less individuated undergoers. Hence, a possible link between undergoer voice in symmetrical voice languages and a transitive clause in languages with transitivity alternations is the presence of an individuated undergoer, separated from the actor of the clause. This observation leads to the following questions: If undergoer voice in a symmetrical voice system is governed by the same factors as a transitive clause in a transitivity-based system, what is the difference between the two systems? Furthermore, if everything boils down to a difference in transitivity, why do symmetrical voice systems exist?

I argue here that a possible answer to these questions concerns the degree to which the participants in a clause are grammaticalized. As became evident in chapter 1, there is a major

difference between the two systems in how information is packaged and presented. Languages with standard transitivity-based systems typically have clear subject and object categories. This was illustrated for Longgu in the examples in (8) in chapter 1. Consider again (8a) and (8b), repeated here as (66a) and (66b).

- Longgu
- (66) a. Mwaa-i      e                  ango  
 snake-SG      3SG.SUBJ      crawl  
 ‘The snake is crawling.’
- b. Mwaa-i      e      ango-vi-a                  vanga                  ngaia  
 snake-SG      3SG.S      crawl-TR-3SG.OBJ      food                  3SG  
 ‘The snake crawled to/for its food.’ (Hill 1992: 58-60)

In these clauses, *mwaa* ‘snake’ is the preverbal S and A arguments in (66a) and (66b) respectively. Both precede the preverbal subject marker *e*, with which they both agree in person and number, namely third person and singular. This subject marker *e* is the same, regardless of whether the preverbal argument is S or A. Hence, S and A together form a morphosyntactic category that is separated from the O argument, which in (66b) is marked on the verb by the suffix *-a*. In other words, the arguments in Longgu clauses are clearly grammaticalized into subject and object relations.

Furthermore, in standard transitivity-based subject/object-systems there is a grammaticalized association between syntactic functions and semantic roles, so that the semantic role that is highest on the so-called thematic hierarchy is mapped to the subject function. In the thematic hierarchy given in (67) the agent is higher or more semantically prominent than the instrument, which in turn is more semantically prominent than the patient/theme.

- (67) Thematic hierarchy (Levin & Rappaport Hovav 2005: 155)  
 Agent > Instrument > Patient/Theme

If the clause in a standard transitivity-based system has an agent, this becomes the subject, otherwise the instrument does, and so on (Levin & Rappaport Hovav 2005: 155-6). Recall the English clauses in (10), repeated here in (68).

- (68) a. The bear ate the fish  
 b. The fish was eaten

In (68a) there is an agent, *the bear*, which then becomes the subject of the clause. In (68b), there is no agent available, so the semantic role that is highest on the thematic hierarchy, which is the patient, *the fish*, becomes subject, and the verb is marked as a passive to indicate a deviance from the default association between the agent role and the subject function. In short, this illustrates how languages with a standard transitivity-based subject/object-system typically have a grammaticalized association between grammatical relations and semantic roles.

A symmetrical voice system, on the other hand, is less grammatically and more pragmatically oriented. As illustrated by the Tagalog examples in (9) in chapter 1, repeated here as (69), a speaker can freely choose which semantic participant he wishes to bring attention to, which is then preceded by the phrase marker *ang*, while the verb is marked by an affix signalling the semantic role of this participant.

- Tagalog
- (69) a. **b-um-ili** ng isda sa tindahan **ang lalake**  
 VC-buy CORE fish OBL store man  
 ‘**The man** bought fish in the store.’
- b. **nbi-bilh-in** ng lalake sa tindahan **ang isda**  
 IRR-buy-VC CORE man OBL store fish  
 ‘The man will buy **the fish** in the store.’
- c. **bi-bilh-an** ng lalake ng isda **ang tindahan**  
 IRR-buy-VC CORE man CORE fish store  
 ‘The man will buy fish in **the store**.’
- d. **ipam-bi-bili** ng lalake ng isda **ang salapi**  
 VC-IRR-buy CORE man CORE fish money  
 ‘The man will buy fish with **the money**.’





receives overt morphological voice marking, as illustrated in (68b). This clearly illustrates that while Tagalog has some form of grammaticalized participant properties, the language does not have clear subject and object relations, like those present in Longgu.

The lack of subject and object relations is even more evident in Äiwoo. While the actor argument and the *ang*-marked argument in Tagalog function as syntactic pivots in certain constructions, this pivot function is completely absent in Äiwoo. In other words, there are no restrictions on the type of argument that can be shared under coordination, no fronting of question words, no raising, no syntactic control and no restrictions on which arguments can launch floating quantifiers (Næss 2015: 288). Concerning relativization, there seems to be a restriction in Äiwoo that the relative clause must correspond to the argument being relativized, so that a relativized actor requires the relative clause to have actor voice, while a relativized undergoer requires the relative clause to have undergoer voice and a relativized circumstantial participant requires the relative clause to be cliticized by =Cä, thus forming circumstantial voice (Næss 2015: 289). Hence, this pattern seems to reflect the one in Tagalog. However, whenever a clause has a relativized actor but also a pronominal undergoer, the relative clause must have undergoer voice. In other words, not even relativization has a consistent restriction on which argument that can be relativized. In short, there is no evidence for the existence of grammatical relations in Äiwoo.

Hence, to answer the first of the questions asked above, a major difference between these two types of systems concerns how or to what degree the participants in a clause are grammaticalized. A standard transitivity-based system has subject and object relations, which are characterized by the grammaticalized association between syntactic functions and semantic roles. In a symmetrical voice system by contrast, the association, if any, is between syntactic functions and pragmatic, rather than semantic prominence. Consequently, this kind of system does not have typical subject and object relations.

Returning to the second question asked above, “Why do symmetrical voice systems exist, if everything boils down to a difference in transitivity?” I argue that an answer to this question must be viewed in light of language change. Looking at the different languages within the Austronesian language family and comparing the clause systems of Proto Austronesian and Proto Malayo-Polynesian with Proto Oceanic, we clearly see the direction the development has taken: from symmetrical voice to transitivity alternations. To the best of my knowledge, no language has developed from a transitivity-based system to a symmetrical voice system. Hence, languages seem to develop from more pragmatic, less grammaticalized

systems into systems with a grammaticalized association between syntactic functions and semantic roles. In other words, this development reflects a transition from a system where syntactic prominence is associated with pragmatic prominence, to a system where syntactic prominence is associated with semantic prominence.

## 5.4 Future research

In this thesis, I have demonstrated how givenness and aspect influence the choice of voice in Äiwoo. Specifically, I have shown how new undergoer arguments and imperfective aspect correlate with actor voice and how given undergoers and perfective aspect correlate with undergoer voice. Based on this evidence, I have argued that undergoer voice in Äiwoo corresponds to a higher degree of semantic transitivity than actor voice, a pattern that is found in a number of other western Austronesian languages with symmetrical voice systems. Together with the facts that UV-clauses are vastly more frequent than AV-clauses, that a number of transitive verbs do not have an attested actor voice form, and that some AV-verb roots are used in UV-constructions even though corresponding UV-verb roots of the same verb are attested, this observation might indicate that undergoer voice will eventually be reanalyzed as the standard transitive construction in Äiwoo and consequently that Äiwoo is in the process of developing into a system of transitivity alternations, like that found in other Oceanic languages.

If one thing has become clear in this thesis, it is that symmetrical voice is a complex phenomenon. As the study by Riesberg et al. (2020) illustrated, however, there are several other factors that can be investigated in order to provide even more answers as to what governs voice choice in Äiwoo, like argument realization and genre or types of text.

The focus of this thesis has been the choice between the two main voices of Äiwoo: actor voice and undergoer voice. As such, I have not considered the motivations behind the choice of the third voice, namely circumstantial voice. Seeing as the opposition between actor voice and undergoer voice seems to concern the degree of individuation of the undergoer and consequently the degree of semantic transitivity, it would be interesting to see how this third clause type patterns with respect to the other two. In other words, because circumstantial voice is formed by the enclitic =Cä which attaches to either an AV-construction or a UV-construction, it is interesting to investigate whether there is some correlation between semantic transitivity and which of these constructions circumstantial voice is formed from.

# References

- Arka, I Wayan and Manning, Christopher D. 2008. Voice and Grammatical Relations in Indonesian: A New Perspective. In *Voice and Grammatical Relations in Austronesian Languages*, Peter K. Austin and Simon Musgrave (eds.) Studies in Constraint-Based Lexicalism.
- Baayen, R. H. 2008. *Analyzing Linguistic Data. A Practical Introduction to Statistics Using R*. Cambridge University Press.
- Beavers, John. 2011. *On affectedness*. Nat Lang Linguistic Theory, 29:335-370.
- Blust, Robert A. 1999. Subgrouping, circularity and extinction: some issues in Austronesian comparative linguistics. In *Selected Papers from the Eighth International Conference on Austronesian Linguistics*, Zeitoun & Li (eds.), 31-94. Taipei: Academia Sinica.
- Chafe, William. 1987. Cognitive Constraints on Information Flow. In *Typological Studies in Language*, Volume 11: Coherence and Grounding in Discourse, Russell S. Tomlin (ed.), 21-51. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Chen, Victoria and McDonnell, Bradley. 2019. *Western Austronesian Voice*. Annual Review of Linguistics, 5:173-95.
- Cumming, Susanna. 1991. *Functional change: The case of Malay constituent order*. Berlin: De Gruyter.
- Du Bois, John W. 1987. *The Discourse Basis of Ergativity*. Language 63 (4), 805-885.
- Foley, William A. 2008. The Place Of Philippine Languages In A Typology Of Voice Systems. In *Voice and Grammatical Relations in Austronesian Languages*, Peter K. Austin and Simon Musgrave (eds.). Studies in Constraint-Based Lexicalism.
- Givón, Talmy. (ed.). 1983. *Topic Continuity in Discourse: A Quantitative Cross-language Study*. Typological Studies in Language 3. John Benjamins Publishing Company.
- Givón, Talmy. (ed.). 1994. *Voice and Inversion*. Typological Studies in Language 28. John Benjamins Publishing Company.
- Hemmings, Charlotte. 2013. *Causatives and Applicatives: The case for Polysemy in Javanese*. SOAS Working Papers in Linguistics, Vol 16.
- Hill, Deborah. 1992. *Longgu grammar*. PhD thesis, Australian National University.
- Himmelman, Nikolaus P. 2002. Voice in western Austronesian: an update. In *The history and typology of western Austronesian voice systems*, Fay Wouk and Malcolm Ross (eds.), 17-62. Canberra: Pacific Linguistics.

- Himmelman, Nikolas P. 2005. The Austronesian languages of Asia and Madagascar: Typological characteristics. In *The Austronesian languages of Asia and Madagascar*, Alexander Adelaar and Nikolaos P. Himmelman (ed.s). Routledge Language Family Series.
- Hopper, P.J. & Thompson, S.A. 1980. *Transitivity in Grammar and Discourse*. *Language*, 56:2, 251– 299.
- Howell, David C. 2007. *Statistical Methods for Psychology*. University of Vermont.
- Jensen, Bård Uri. 2018. Er resultatet gyldig? Noen utfordringer ved bruk av kvantitative metoder i andrespråksforskning. In *Norsk som andrespråk – perspektiver på læring og utvikling*, Ann-Kristin Helland Gujord and Gunhild Tveit Randen (eds.).
- Kemmer, Suzanne. 1993. *The Middle Voice*. Typological Studies in Language 23. John Benjamins Publishing Company.
- Kroeger, Paul. 1993. *Phrase Structure and Grammatical Relations in Tagalog*. Dissertation, Center for the Study of Language and Information.
- Levin, Beth and Rappaport Hovav, Malka. 2005. *Argument Realization*. Cambridge University Press.
- Liao, Hsiu-chuan. 2004. *Transitivity and Ergativity in Formosan and Philippine Languages*. University of Hawai'i.
- Lynch, John; Ross, Malcom and Crowley, Terry (eds.). 2011 *The Oceanic Languages*. Routledge Language Family Series.
- Margetts, Anna. 2008. *Transitivity Discord in some Oceanic Languages*. *Oceanic Linguistics*, Volume 47, no. 1. University of Hawai'i Press. Monash University.
- McDonnell, Bradley James. 2016. *Symmetrical Voice Constructions in Besemah: A Usage-based Approach*. University of California, Santa Barbara, dissertation.
- Musgrave, Simon. 2008. Introduction. In *Voice and Grammatical Relations in Austronesian Languages*, Peter K. Austin and Simon Musgrave (eds.). *Studies in Constraint-Based Lexicalism*.
- Næss, Åshild. 2007. *Prototypical Transitivity*. Typological Studies in Language 72. John Benjamins Publishing Company.
- Næss, Åshild. 2015. *Voice at the Crossroads: Symmetrical Clause Alternations in Äiwoo, Reef Islands, Solomon Islands*. *Oceanic Linguistics* 54:1, 270-307.
- Pastika, I Wayan. 1999. *Voice Selection in Balinese Discourse*. Canberra: Australian National University dissertation.

- Quick, Phil. 2005. Topic Continuity, Voice and Word Order in Pendau. In *The many faces of Austronesian voice systems: some new empirical studies*, Wayan Arka and Malcolm Ross, (eds.), 221-242.
- Rietveld, Toni, van Hout, Roeland and Ernestus, Mirjam. 2004. *Pitfalls in Corpus Research*. Computers and the Humanities 38: 343-362. Kluwer Academic Publishers.
- Rietveld, Toni and van Hout, Roeland. 2005. *Statistics in Language Research. Analysis of Variance*. De Gruyter Mouton.
- Riesberg, Sonja. 2014. *Symmetrical Voice and Linking in Western Austronesian Language*. De Gruyter Mouton.
- Riesberg, Sonja; Malcher, Kurt; Bardají I Farré, Maria and Himmelmann, Nikolaus P. 2020. *Voice choice in Totoli discourse*. Paper presented at the 12<sup>th</sup> International Austronesian and Papuan Languages and Linguistics Conference, Oslo June 2020.
- Ross, Malcolm D. 2002. The history and transitivity of western Austronesian voice and voice-marking. In *The history and typology of western Austronesian voice systems*, Fay Wouk and Malcolm Ross (eds.), 17-62. Canberra: Pacific Linguistics.
- Ross, Malcolm D. 2004. *The Morphosyntactic Typology of Oceanic Languages*. Language and Linguistics 5.2:491-541. The Australian National University.
- Ross, Malcolm D. and Næss, Åshild. 2007. *An Oceanic Origin for Äiwoo, the Language of the Reef Islands?* Oceanic Linguistics, Volume 46, no. 2. University of Hawai'i Press.
- Roversi, Giovanni. 2019. *The morphosyntactic structure of the Äiwoo verb*. University of Oslo.
- Roversi, Giovanni and Næss, Åshild. 2019. *Jespersen in the Reef Islands: Single versus Bipartite Negation in Äiwoo*. Oceanic Linguistics, 58:2, 325-51.
- Saeed, John I. 2016. *Semantics*. Fourth Edition. Wiley Blackwell.
- Schachter, Paul. 1977. Reference-related and role-related properties of subjects. In *Syntax and Semantics 8. Grammatical Relations*, 279-306.
- Van Valin, Robert D. Jr and LaPolla, Randy J. 1997. *Syntax. Structure, meaning and function*. Cambridge University Press.
- Walters, Dennis. 1994. Discourse-Based Evidence for an Ergative Analysis of Cebuano. In *UTA Working Papers in Linguistics 1*, Susan C. Herring & John C. Paolillo, (eds.).
- Wouk, Fay. 1996. *Voice in Indonesian discourse and its implications for theories of the development of ergativity*. Studies in Language 20:2, 361-410. University of Auckland.

Wouk, Fay. 1999. *Sasak is Different: A Discourse Perspective on Voice*. Oceanic Linguistics, Volume 38. University of Hawai'i Press.