# Events in Language and Thought:
## The Case of Serial Verb Constructions in Avatime

Printed and bound by Ipskamp Drukkers, Nijmegen
Cover photo: Goats eating decorations after Holy Cross celebrations in Vane, Avatime, 2011

# Events in Language and Thought:
## The Case of Serial Verb Constructions in Avatime

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken,
volgens besluit van het college van decanen
in het openbaar te verdedigen op donderdag 3 november 2016
om 10.30 uur precies

door
Rebecca Jane Defina

geboren op 15 april 1983
te Sydney, Australië

*Promotoren:*
Prof. dr. Stephen C. Levinson
Prof. dr. Asifa Majid

*Copromotor:*
Dr. Felix K. Ameka (Universiteit Leiden)

*Manuscriptcommissie:*
Prof. dr. Aslı Özyürek
Prof. dr. Birgit Hellwig (Universität zu Köln, Duitsland)
Prof. dr. Jeffrey M. Zacks (Washington University in St. Louis, Verenigde Staten)

The research reported in this thesis was supported by the Max-Planck-Gesellschaft zur Förderung der Wissenschaften, München, Germany.

# Events in Language and Thought:
## The Case of Serial Verb Constructions in Avatime

Doctoral Thesis

to obtain the degree of doctor
from Radboud University Nijmegen
on the authority of the Rector Magnificus prof. dr. J.H.J.M. van Krieken,
according to the decision of the Council of Deans
to be defended in public on Thursday, November 3, 2016
at 10.30 hours

by
Rebecca Jane Defina

Born on 15 April 1983
in Sydney, Australia

*Supervisors:*
Prof. dr. Stephen C. Levinson
Prof. dr. Asifa Majid

*Co-supervisor:*
Dr. Felix K. Ameka (Leiden University)

*Doctoral Thesis Committee:*
Prof. dr. Aslı Özyürek
Prof. dr. Birgit Hellwig (University of Cologne, Germany)
Prof. dr. Jeffrey M. Zacks (Washington University in St. Louis, United States of America)

The research reported in this thesis was supported by the Max-Planck-Gesellschaft zur Förderung der Wissenschaften, München, Germany.

# Contents

## Acknowledgements

This thesis would not exist were it not for the help and kindness of many people. I would like to thank everyone who has assisted me along the way.

Firstly, I would like to thank the Avatime people without whose assistance this project would have been impossible. Thank you to the Paramount Chief Osie Adza Tekpor VII, Paramount Queen Mother Onetsitsie Osei Yawa VIII and the chiefs and elders of Vane for your warm welcomes and continuing support. For assistance with the various logistical aspects of fieldwork I thank Charles Gbagbo, Wisdom Ekissi, Sacha and Lena Vinokurov, and Divine Mununkum. I'd also like to thank Divine for his assistance in filming the videos used to create the stimuli items in Chapter 5. On the wall of the Junior High School in Gbadzeme someone has written "JHS has come to save all". While I'm sure it was by no means the original intention, these words reflected my own experience as well. The generous hospitality and assistance of the Amedzofe, Biakpa, Gbadzeme and Vane Junior High Schools really saved this research project. Thank you to all the principals, teachers and students for making me so welcome. Finally, I am very grateful to all the Avatime people who allowed me to record them, answered my questions, or assisted with transcription and analysis. Five people have been especially helpful as consultants and friends: Delali Quansah[1], Prosper Dugboryele, Mathias Mahunu, Sammy Oboni, and my Avatime mother Charlotte Adzoyo Bakudie. *Mlɔ mlɛwa lixwɛ popopo!*

I am immensely grateful to my supervisors Stephen Levinson, Felix Ameka and Asifa Majid for all their support and assistance. Steve, it has been a real pleasure and an incredibly enlightening experience to work with you and the L&C. I have greatly valued our discussions and all your thorough, kind, illuminating, and challenging feedback, thank you. Felix, I am very grateful for your continued support throughout both my MPhil and PhD. Your kind and thoughtful advice has been invaluable throughout my work on Avatime. I have learned a lot from our discussions/debates and I look forward to more in the future. Asifa, having you as my day-to-day supervisor has been an amazing experience, thank you. You have given very generously of your time not only to supervise this thesis but to guide and train me in developing as a scientist. I have learnt so much from your patient yet persistent advice and detailed comments. I've also really appreciated the support

[1] who sadly passed away much too soon and is still greatly missed

you've given me and the times you've helped put things in perspective. I was briefly tempted by your offer to let me submit my thesis through interpretative dance, but the horror that came with the idea of starting back at the beginning made me realise how much progress I'd actually already made. I sometimes wonder how it would have gone if I'd tried to take you up on that though. All three of my supervisors have provided outstanding role models as scientists, academics, and supervisors and I hope to be able to live up to their examples.

I would also like to think my manuscript committee – Aslı Özyürek, Birgit Hellwig and Jeff Zacks – and all the corona members for taking the time to read and evaluate this thesis.

Doing this PhD at the MPI in Nijmegen has been an amazing experience both in terms of the support I received and the colleagues I got the chance to work with and talk to. Thank you to Edith Sjoerdsma for all your assistance with administrative matters and your always friendly conversations which never failed to lighten up my day, especially towards the end. Thanks to the IMPRS team – Rachel Sheer, Els den Os, Dirkje van den Aa and Kevin Lam – for all your administrative and moral support. Thanks also to the library staff, TG, and digiteam for all your assistance. Especially to Ronald Fisher for his assistance with writing the Presentation scripts to run my experiments, including phone support while I was in Ghana, and to Nick Wood for his ever speedy and excellent assistance with data conversions. Thanks to all those who acted in my stimuli videos. There was one day where I believe I managed to hassle everyone who happened to be in the building so I won't list you all here, but thank you! Thanks also to Emiel van den Hoven and Frédérique Schless for assisting with the testing of Dutch participants for the experiments reported in Chapters 5 and 6.

For stimulating discussions, fun times, and copious support I would like to thank my various office mates Tyko Dirksmeyer, Gabriela Garrido, Saskia van Putten, and Ewelina Wnuk as well as my other fellow PhD students and colleagues: Inge Alferink, Julija Baranova, Joe Blythe, Marisa Casillas, Sarah Cutfield, Mark Dingemanse, Sarah Dolscheid, Michael Dunn, Nick Enfield, Simeon Floyd, Reyhan Furman, Binyam Gebrekidan Gebre, Rósa Signý Gisladottir, Harold Hammarström, Jeremy Hammond, Franziska Hartung, Kaoru Hayano, Clair Hill, Elliot Hoey, Judith Holler, Gertie Hoyman, Kobin Kendrick, Vishnupriya Kolipakam, Dejan Matić, Elisabeth Manrique, Elisabeth Norcliffe, Sean Roberts, Giovanni Rossi, Lila San Roque, Sebastian Sauppe, Gunter Senft, Hilario de Sousa, Robert Van Valin, Maartje van de

Velde, Annemarie Verkerk, Connie de Vos, Marine Vuillermet, Peter Withers, and others. I was also very fortunate to meet and befriend the late Melissa Bowerman; I have fond memories of our discussions and many of the questions and comments you raised still linger with me.

Some of these people made substantial contributions and cannot get off so lightly without additional thanks. Thank you Julija for our babysitting/working sessions. They provided equally well appreciated work time and conversation. Thank you Marisa for helping me get started with linear mixed effects modeling. The study reported in Chapter 4 began life as a research project within Aslı Özyürek's Gesture and Sign course. Thank you Aslı, for your inspiration and guidance. Special thanks also go to Binyam for attempting to automatically annotate the gestures in my data and assisting with the comparison between dwell times and pixel changes in Chapter 5. Thank you to Saskia for translating my samenvatting. Finally, thank you also to all those who read and commented on parts of this thesis: our incredibly supportive writing group – Ewelina, Rósa, Saskia, Julija, Ely and Gaby, as well as Asli Özyürek, Gunter Senft, Robert Van Valin, and Anelie Crighton.

My move to the Netherlands to work on African languages at Leiden University was a pivot point in my career. Thank you to Maarten Mous for chatting with me in Australia and inspiring and encouraging me to make the move.

Thank you to all my friends and family for your encouragement and support. To Anelie for reminding me that the only way to eat an elephant is one bite at a time. To Thalie for providing me some insight into what doing a PhD could be like and hope that I might also one day finish it. To Fannie for introducing me to electro swing which kept me going during the final stages of writing and revising. To my parents, Wendi and Greg, for their unwavering support.

I feel honoured to have the outstanding linguists and my fantastic friends Saskia van Putten and Ewelina Wnuk standing with me as my paranimfen. Saskia, it has been amazing working with you over the last 9 years – doing fieldwork together, sharing our data and working out some of the puzzles of Avatime. Thank you for all of the discussions, fun times, and assistance in times of need. Ewelina, I am so glad we got to meet and journey through our PhDs together. Your understanding and insight have been incredibly helpful on many occasions, thank you for everything.

And finally to my husband Brahn and our children Theo and Max. Theo, I started this PhD shortly after you were born and have been working on it for your whole life so far. Thank you for motivating me to improve my time management skills and

for helping me put this endeavour in perspective. Your growth and development have provided me with a very real and tangible timeline to use as reference for my own progress, although your accomplishments often seemed to eclipse my own. Max, we shared the final stages of writing and revising this thesis. Thank you for sleeping so well and for brightening my days. Finally, Brahn, thank you for accompanying me on this wacky adventure, for your unflappable support and confidence, and your extensive practical assistance especially during field trips. I'm looking forward to many more good times with you all.

## Abbreviations

| | |
|---|---|
| 1 | ‘first person’ |
| 2 | ‘second person’ |
| ADD | ‘additive’ |
| ATR | ‘Advanced tongue root’ |
| $C_X$ | ‘noun class x’ |
| CERT | ‘certainty’ |
| CLM | ‘clause linkage marker’ |
| CM | ‘clause marker’ |
| COM | ‘comitative’ |
| COMP | ‘complementizer’ |
| CTR | ‘contrastive’ |
| DEF | ‘definite’ |
| DIST | ‘distal’ |
| FOC | ‘focus’ |
| FP | ‘final particle’ |
| GILLBT | ‘Ghana Institute for Linguistics, Literacy and Bible Translation’ |
| GTM | ‘Ghana-Togo Mountain’ |
| HAB | ‘habitual’ |
| ID | ‘ideophone’ |
| INDEF | ‘indefinite’ |
| INF | ‘infinitive’ |
| INT | ‘intentive’ |
| IT | ‘itive’ |
| LD | ‘left dislocated’ |
| LOC | ‘locative’ |
| LOG | ‘logophoric’ |
| NEG | ‘negative’ |
| OBJ | ‘object’ |
| p | ‘plural’ |
| PFV | ‘perfective’ |
| POS | ‘possessive’ |
| POSM | ‘possessum’ |
| POT | ‘potential’ |
| PROG | ‘progressive’ |
| PROH | ‘prohibitive’ |
| PROX | ‘proximal’ |
| REC | ‘recurrent’ |
| REDUP | ‘reduplication’ |
| REL | ‘relative clause marker’ |
| s | ‘singular’ |
| SBJ | ‘subject’ |
| SBJV | ‘subjunctive’ |
| SVC | ‘serial verb construction’ |
| SVM | ‘serial verb construction reduced agreement marker’ |
| VENT | ‘ventive’ |

# 1 Introduction

> Events are not simply out there and ready-made, waiting to be seen, recognized, or described; they are what we make of them.
>
> (Schwartz 2008:54)

> It is becoming clear that the semantic domain which is most 'miscellaneous and gerrymandered' is that of event terms… The lexicalization of event-denoting expressions has thus emerged as the domain with the most extreme cross-linguistic variation. This makes the mapping between the semantic level of 'event' and the syntactic level of 'clause' one of the greatest challenges to semantics, syntax and typology. (Evans 2010:5)

Events occur all around us in continuous streams of activity, yet we think and talk about them in terms of discrete units. These units are not pre-determined by natural boundaries, but rather, we – as observers, cognizers, and describers – construct them (Schwartz 2008). As constructed entities, they are potentially subject to variation across individual observers and contexts. In fact, studies of event segmentation during perception routinely note considerable individual differences (e.g. Zacks, Tversky & Iyer 2001). These studies have all focused on speakers of English. However, as Evans (2010) notes, there is extreme variation in how different languages encode events. The way we segment events in language is often assumed to mirror the way we segment them conceptually (e.g. Davidson 1969; Folli & Harley 2006; Malaia 2014). If true, this would imply people who speak different languages vary greatly in terms of how they segment events conceptually. In this thesis, I examine the nature and extent of the alignment between conceptual and linguistic event units and discuss the implications for the ways in which we discretize activity.

For this study, I focus on one particular language – Avatime, a Kwa language spoken in Ghana – and a particular type of syntactic construction – serial verb constructions. These are constructions where multiple verbs are combined within a single clause, much like the English *Let's go grab a drink*. They provide an excellent opportunity for investigating issues of event segmentation and the alignment between linguistic and conceptual units. The whole construction consists of a single clause and is often claimed to refer to a single conceptual event unit (e.g.

Aikhenvald 2006; Bisang 2009; Comrie 1995). However, each verb provides a semantic contribution to the overall event description and could independently head its own predicate. They are, thus, explicit in drawing event information from multiple sources within the construction, and some have suggested they are better treated as collections of sub-events (e.g. Baker & Harvey 2010; Pawley 2011). Although serial verb constructions occur frequently among the world's languages, they are rare among the more often discussed European languages (Dixon 2006).

This investigation of serial verb constructions within Avatime employs a variety of methods to target different aspects of linguistic and conceptual event segmentation, including linguistic description, gesture analysis, and behavioural experiments in which Avatime speakers are compared to speakers of Dutch and English. This multi-method approach, while focusing in on one type of syntactic construction in one language, allows for a detailed investigation of the various relationships between linguistic and conceptual event structures. In the remainder of this introduction, I describe events and serial verb constructions in more detail, discuss previous work on the relations between events in language and thought, and finally outline the specific methods and goals of this thesis.

## 1.1 What is an event?

The discussion of what events are, and how to identify and distinguish them, has a very long history and is still very far from reaching a resolution (see Casati & Varzi 1996; Casati & Varzi 2008 for excellent overviews). I do not presume to suggest the issues involved will be resolved here. Rather, my goal is to describe the issues relevant for the present investigation and explicate the position taken in this thesis and the assumptions made.

First, a distinction must be drawn between event tokens and types. Tokens are particular instances which are never repeated, such as the particular time you brushed your teeth on the morning of October 13th 2014, or you reading this thesis right now. Types group together particular event tokens into kinds of events, for instance tooth brushing or reading. Particular tokens of these event types are never exactly the same, for instance even the second reading of the same text while sitting in the same place will be a different experience to the first. Nevertheless, all acts of reading share sufficient commonalities for us to group them together as instances of the same event type. In this thesis, my primary focus is on the representation of

individual event tokens and how particular instances are discretized during perception, memory, and description. This focus on particular event tokens allows a grounding of reference in the specific real world occurrences which facilitates comparison across individuals as well as between levels of representation.

Secondly, there is the question of what kinds of things count as an event. Events have been divided into several different types or Aktionsarten. This division was first documented in the work of Aristotle in *Metaphysics* 1048b and has been further developed more recently by Ryle (1949), Kenny (1963), and Vendler (1957) among others. Vendler's classification is the most commonly used and distinguishes four types: states, activities, accomplishments, and achievements. States are characterized by a lack of change over (a period of) time (e.g. *being tired*). Activities are characterized by dynamic action without an inherent end-point (e.g. *running*). In contrast, accomplishments do have an inherent end-point (e.g. *writing a thesis*). Achievements are characterized by an instantaneous change of state (e.g. the eureka moment of *discovery*). While these distinctions have had a massive impact on the study of events in philosophy and linguistics (see Sasse 2002 for further discussion), the main concern for the present thesis is that researchers differ regarding which of these types they include as 'events'. While accomplishments and achievements are generally agreed to be events, some draw the line there and do not include states or activities as events (e.g. Bach 1986; Mourelatos 1978), others include actions (e.g. Smith 1999), and yet others include all four types (e.g. Rothstein 2005). In this thesis, I take the more expansive view of events including states and actions. This inclusion is of most relevance to the co-speech gesture study in Chapter 4 where a wide variety of events are included. The studies in Chapters 5 and 6 focus on more typical, accomplishment-type events.

The other major issue regarding *what* an event is relates to *where* it is. Von Stutterheim and Nuse (2003) clearly articulated this issue when they drew the following distinctions between events:

- as they occur in the real world (here referred to as actions, activities, or states)
- as we conceptualize them (here referred to as events)
- and as we describe them (here referred to as event descriptions)

This connects with Schwartz's (2008) account of events as constructed entities. It also highlights the fact that, while the events we create in our linguistic descriptions refer and relate to our conceptual representations and the real world occurrences,

they are unlikely to mirror them exactly. Furthermore, it is worthwhile making even finer distinctions and dividing conceptual events into at least the following:

- the events we create during perception
- those we store in memory
- and the representations we form for the purposes of description and communication.

These distinctions set up the architecture of what is included as an event and what kinds of event representations are considered within this thesis, but leaves us with two fundamental questions: how do we know whether two event representations or descriptions relate to the same event? And what does it mean for something to be conceived of as *one* event?

The question of how to judge whether two conceptual event representations or descriptions (since the two are often conflated) refer to the same event unit or not has plagued philosophers for a long time (e.g. Cleland 1991; Davidson 1969; Goldman 1971). For instance, is *flipping the light switch* the same as *turning on the light*, or is it a subcomponent of that event? There have been three main approaches to resolving these questions, each focusing on different criteria: 1) whether the events have the same causes and effects (Davidson 1969); 2) whether they occupy the same portion of space and time (Quine 1985); and 3) whether they are exemplifications by the same objects of the same properties at the same times (Kim 1976). The cause and effect approach of Davidson has generally received wider acceptance though all three are still active (Goldman 2007). It is in fact likely that all three types of criteria influence our ideas of whether or not two event representations reference the same event. For the purposes of the present thesis, I assume coreference of event representations so long as they meet any of the above criteria.

The question of whether something is a single event or a collection of multiple events intersects with the above discussions in philosophy on how to identify events, for if *flipping the light switch* is not the same as *turning on the light* then it must be a subevent of the more complex *turning on the light* event. Approaches like Davidson's and Quine's tend to lead to more unified event representations, whereas approaches like Kim's tend to favour division into finer-grained, smaller event units (Casati & Varzi 1996). Event segmentation has also been a major focus of the psychological investigation of event perception and cognition (e.g. Newtson & Engquist 1976;

Shipley & Zacks 2008). Within psychology, there is currently one clearly dominant theory: the Event Segmentation Theory proposed by Zacks and colleagues (Zacks et al. 2007). According to this theory, current perceptual information is combined with information from episodic memory and event schemata (c.f. frames Fillmore 1977; and scripts Schank & Abelson 1977) to form a model of the current event and so predict upcoming perceptual information. If the event model no longer makes accurate predictions, it is updated. The times when the event model is updated correspond to the times when one event is perceived to finish and another begins, i.e. to the boundaries between events. Event segmentation is thus influenced by information from the perceptual input, such as properties of the actor's motion, changes in location, or changes of object, and by information contained in the active event schemata, such as expectations regarding goals and typical ways of achieving these goals. It is these event schemata which may mediate an influence of language on event segmentation, since language may influence the contents of specific event schemata or which schemata are activated at different times.

One of the main findings of the psychological research on event segmentation is that events are segmented in a hierarchical manner (Zacks, Tversky & Iyer 2001): they consist of subevents and are themselves parts of larger event units. For instance *conducting a study* consists of many subevents of *designing the experiment*, *preparing the materials*, *collecting the data, analysing the results*, etc., and it is also a part of the larger event of *doing a PhD*. People are aware of these different levels of event segmentation and can segment events at fine or coarse levels depending on the requirements of the context (e.g. Newtson 1973; Zacks, Tversky & Iyer 2001). These fine and coarse level boundaries align with each other (Zacks, Tversky & Iyer 2001), suggesting both segmentations are driven by a single hierarchical event representation. This means there is an essential duality whereby event units can be viewed as whole units but also as complex entities made up of their constituent subevents.

This parallels the situation with objects which also have hierarchical partonymic structures (e.g. all the various parts of the human body and their subcomponents). Objects are also often classified in hierarchical taxonomic structures (e.g. a teaspoon is a kind of spoon). Similar taxonomic structures can be seen among event types (Tversky 1990; Zacks & Tversky 2001). Research on objects has suggested there is a basic level in this taxonomy which is most commonly used when describing things (Rosch et al. 1976; Rosch, Simpson & Miller 1976). Some attempts have been made

to find a corresponding basic level for events. Rifkin (1985) and Morris and Murphy (1990) investigated event taxonomies using nominal event descriptions such as *meal*, *breakfast*, and *quick breakfast* or *entertainment, the movies,* and *horror movie*. They both found evidence for a basic level but it wasn't as stable as the basic level of objects. For instance, the tendency to refer to an event with a basic level rather than subordinate term depended on context (Morris & Murphy 1990) contrary to what Rosch et al. (1976) found for object names. The use of nominal event descriptions also made it unclear whether participants were really judging events. In any case, the study of a basic level for events is much less developed than that of objects.

The philosophical and psychological investigations of events have overwhelmingly been carried out in English. The ways in which events are encoded in language are, however, extremely variable (e.g. Evans 2010; Gentner 1982; Malt & Majid 2013; Talmy 1985; Talmy 2000). Languages differ in terms of which event types they lexicalize in verbs, how they encode the participants of these verbs, and what kinds of syntactic structures they employ (e.g. Bohnemeyer & Pederson 2011; Bowerman & Brown 2008; Talmy 2000). This means previous investigations into event concepts may have been influenced by peculiar patterns of how English speakers represent events. The following section reviews previous work addressing the possible influence of language on thought in general and on event conceptualizations in particular.

## 1.2 Relations between language and thought

Debate regarding linguistic relativity – whether language influences thought – has been long and turbulent (see Bloom & Keil 2001; and Levinson 2012 for excellent overviews). The topic has triggered strong reactions among many researchers, and sweeping statements in both directions, from "Language is identical with thought" (Müller 1909:ii) to "But it is wrong, all wrong" (Pinker 1994:57). Recently, there has been increased focus on distinguishing the different ways language may influence thought. These finer distinctions are helping form a clearer understanding of the complex relationships between language and thought, and moving away from the overly simplified, polarised yes or no debate (e.g. Bloom & Keil 2001; Wolff & Holmes 2011).

There are several ways language clearly and uncontroversially influences thought, after all "this is what language is *for*" (Bloom & Keil 2001:354). We use it

to gossip with friends, teach our children, and argue with ourselves. It is a powerful tool for the expression and transfer of ideas and much of our mental content has been delivered and refined by language. The question which has bothered so many is whether language influences thought in some *other* way, not just through the content it communicates. Unfortunately, it is often very difficult to isolate such an influence from both semantic content and any language-external factors.

One well-established instance of linguistic influence on the way we think occurs *during* language use. This influence is commonly referred to as 'thinking-for-speaking' (Slobin 1987; Slobin 1996a). It also has a long history in work on sentence production (e.g. Levelt 1989). In order to produce a sentence, one must form a 'conceptual message' containing the information one wishes to express. The content of this message and the process used to produce it are influenced by the properties of the language used (e.g. Bock 1982; Chang, Dell & Bock 2006; Kita & Özyürek 2003; Levelt 1989; Papafragou, Hulbert & Trueswell 2008). For instance, Kita and Özyürek (2003) found differences in the content of the conceptual messages of English, Japanese, and Turkish speakers describing a cartoon showing the cat Sylvester swinging from one building to another. While English has an easily accessible way of encoding agentive motion along an arc – the verb 'swing' – Japanese and Turkish do not. Correspondingly, English speakers frequently produced arc shaped gestures when describing this scene, while Japanese and Turkish speakers tended to produce straight gestures. This suggests the arc trajectory was part of the conceptual message for English speakers but abstracted away by Japanese and Turkish speakers. While studies such as this suggest an influence of linguistic structure on the conceptual message, it is worth noting that the conceptual message is not perfectly aligned with the linguistic structure. For instance, the left-right direction of motion is rarely reported in speech, but participants in Kita and Özyürek's study regularly reproduced the direction in their gestures. Thus, the semantic content of speech is closely related, but not isomorphic, to the conceptual message.

The influence of active language use can also be seen in the way people solve problems or remember things. For instance, the way people discriminate colours can be influenced by the language they use (e.g. Gilbert, Regier & Kay 2006; Roberson, Davies & Davidoff 2000; Winawer et al. 2007). People are faster at distinguishing colours which their language places into distinct colour categories. This effect is restricted under verbal interference (Winawer et al. 2007) and is stronger when the

colours are presented in the right visual field (Gilbert, Regier & Kay 2006). Thus the effects appear to be driven by the active use of language. This has led some to discount these effects, claiming that *real* linguistic relativity effects should not be mediated by the active use of language (e.g. Papafragou, Massey & Gleitman 2002).

More lasting influences of habitual language use have also been noted in various domains such as: gender assignment (Boroditsky, Schmidt & Phillips 2003; Vigliocco et al. 2005), number (Frank et al. 2008), sound (Dolscheid et al. 2013), spatial frames of reference (Majid et al. 2004), spatial relations (Choi 2006; McDonough, Choi & Mandler 2003), and time (Boroditsky 2011). For instance, Korean makes a distinction between tight and loose fitting containment and this leads Korean speakers to focus more on this distinction than English speakers (Choi 2006; McDonough, Choi & Mandler 2003). Choi (2006) further showed English and Korean infants attend equally to the tight versus loose distinction, but English speaking children's sensitivity is reduced by the age of 3 years. While there is a lasting difference in the amount of attention given to the tightness of the fit, this does not mean English speakers cease to distinguish between tight and loose fit. After all, we still need to get on with our daily lives and as Malt (2012) noted: the language you speak is not likely to change how tightly you put the nappy on a baby. Thus, while habitual language use can influence the way we think, personal experience and specific goal requirements remain major factors (Dougherty & Keller 1982; Malt et al. 1999).

In review, variation in the structure of languages and the way they encode information can influence the way we think in various ways. This influence can neither be completely discarded as Pinker (1994) claimed, nor is it as complete and deterministic as has been suggested by some other authors (e.g. Müller 1909). In the following section, I review in more detail the previous research specifically focusing on the relations between language and cognition in regard to events.

## 1.3 Events in language and thought

First of all, active language use and the variation found within even a single language have been shown to influence the way events are segmented and remembered. The mere use of language to describe events while people are segmenting them has been shown to lead people to segment in a more hierarchical manner with greater agreement between segmentations over different viewings

(Zacks et al. 2001). The particular characteristics of the language used have also been shown to have an influence. For instance, using imperfective or progressive, as opposed to perfective, aspect appears to lead people to segment events more finely (Matlock et al. 2012). The particular semantic characteristics of the verbs used can also have an influence on how events are remembered. In their classic study, Loftus and Palmer (1974) showed people videos of car crashes and asked them how fast the cars were going when they "smashed/collided/bumped/hit/contacted". The particular verb used influenced the speeds participants reported in their accounts of the crashes they had seen.

The ways events are described between different languages vary greatly. Indeed, Evans describes events as "the domain with the most extreme cross-linguistic variation" (Evans 2010:5). According to Gentner (Gentner 1982; Gentner 2006; Gentner & Boroditsky 2001), there are two motivations for this great variation: 1) events relate to less cohesive collections of percepts than objects; and 2) there are more degrees of freedom for lexicalizing events, which involve relations between objects, than in lexicalizing objects themselves. Both of these motivations suggest the variation in how events are described may extend to variation in how they are conceptualized.

This makes events a prime area for investigating linguistic relativity. However, variation in how speakers of different languages think about events has received comparably little research. Most studies have focused on only one distinction: verb- versus satellite-framed languages (Talmy 1985; Talmy 2000). Verb-framed languages tend to encode path information in verbs (e.g. *traverse)* with manner information optional and typically encoded in adverbials (e.g. *wobblingly*). Satellite-framed languages, on the other hand, tend to encode path in satellites like *across*, while manner is encoded in the main verb (e.g. *wobble*). Correspondingly, satellite-framed languages are said to encode manner more frequently and specifically than verb-framed languages (Slobin 1996a; Slobin 1996b). This has led to the prediction that speakers of satellite-framed languages will attend more to manner than speakers of verb-framed languages. This prediction has been tested in several studies using methods such as: similarity judgements (e.g. Bohnemeyer, Eisenbeiss & Narasimhan 2006; Finkbeiner et al. 2002; Gennari et al. 2002; Loucks & Pederson 2011; Papafragou, Massey & Gleitman 2002), recognition memory (e.g. Filipović 2011; Gennari et al. 2002), novel word learning (e.g. Kersten et al. 2010), and eye tracking (e.g. Papafragou, Hulbert & Trueswell 2008; Trueswell & Papafragou 2010).

The results have been mixed and overall these studies suggest a few fundamental problems.

Firstly, there are questions regarding the prediction that manner will be more salient for speakers of satellite-framed languages. While initial results with English and Spanish suggested satellite-framed languages made more manner distinctions than verb-framed languages (Slobin 1996b), subsequent work suggests the number of manner distinctions is not closely correlated with the verb- versus satellite-framed distinction (Matsumoto 2003). It is also not clear whether encoding manner information in the main verb makes it more salient. In fact, Papafragou and colleagues (Papafragou, Hulbert & Trueswell 2008) found participants looked longer at information which was not encoded in the main verb.

Secondly, initial work took the distinction between verb- and satellite-framed languages as absolute. However, increasing evidence shows this is far from the case (e.g. Beavers, Levin & Tham 2010; Croft et al. 2010; Slobin 2004; Verkerk 2014). While some languages make more use of one strategy than another, languages generally employ both strategies to some extent. Verb- or satellite-framing is then a relative ranking of the proportion of construction usage rather than an absolute typological distinction. This variability in encoding strategies heightens the influence of task-specific effects. For instance, Loucks and Pederson (2011) found that in some tasks both English and Japanese speakers were strongly biased to describe and attend to manner, with no group differences in linguistic framing strategies or non-linguistic task performance.

Finally, it seems manner and path may not, in fact, be well-suited for event cognition research. According to Loucks and Pederson they may even be the two components of motion events "*least likely* to evidence categorizational memory, or attentional differences across linguistically-defined populations" (Loucks & Pederson 2011:133). This is because manner and path are not perceptually distinct, since perceiving one necessarily involves perceiving the other. Path is also basic to all motion event descriptions and is regularly encoded in all languages, either as a verb or satellite, which is why Talmy defined his typology with respect to where the path information was encoded.

More recently, there has been some linguistic relativity research focusing on causal events (e.g. Fausey & Boroditsky 2011; Fausey et al. 2010; Le Guen et al. 2015; Wolff 2003; Wolff & Ventura 2009; Wolff, Jeon & Li 2009). Wolff et al.'s

(2009) study is particularly relevant for this thesis as it concerns the segmentation of events. Wolff et al. investigated the influence of whether a language allows inanimates to function as causers (as in *the key opened the door*) on how people segment chains of causal events. While some languages, such as English and Mandarin Chinese, allow inanimates to function as causers, other languages, such as Korean do not. Wolff and colleagues found this difference does influence the way people segment causal chains with inanimate intermediaries, such as when a girl pops a balloon by placing it under a hot lamp. English and Chinese speakers were more likely to treat such causal chains as multiple events, while Korean speakers were more likely to view them as a single event. This shows an influence of language structure on conceptual event segmentation.

These previous studies have examined variations in syntactic packaging which are specific to certain semantic domains and often intimately connected with semantic variations. For instance, the distinction tested by Wolff et al. (2009) stemmed from a difference in whether or not the language allowed an inanimate intermediary to function as a causer. In contrast, this thesis considers one kind of syntactic construction – serial verb constructions – across a range of semantic domains. This enables a more focused investigation of the relations between syntactic and conceptual event segmentation.

## 1.4 What are serial verb constructions?

Serial verb constructions (SVCs) (such as that in (1)) were first described by Christaller in his 1875 grammar of Twi (1875:69–73). Since then they have been documented in many other West African languages, as well as other regions of the world such as Oceania, South-east Asia, and Amazonia. They are estimated to occur in one-third of the world's languages (Dixon 2006), though they are not evenly distributed. They are ubiquitous in some areas, such as West Africa, and seemingly absent in others, such as Europe.

(1) *bíà-dzì* *va* *ní* *ke-dà*
$C_1$p.POT-return come LOC $C_6$s-back
'They will come back again.' (Avatime-history_110905_BB_134)

Defining SVCs has been a "sticky business" (Lord 1993:1). Many definitions have been proposed over the last few decades (e.g. Aikhenvald 2006; Comrie 1995;

Crowley 2002; Durie 1997; Foley & Olson 1985; Haspelmath 2016; Noonan 1985). The list of criteria proposed by Durie (1997) then adopted by Aikhenvald (2006) and many others, is probably the most exhaustive, containing most of the elements proposed in earlier definitions. According to this list, a SVC should have the following properties:

1. A sequence of two or more verbs which can both also function independently
2. Monoclausality, with all the intonational properties of a monoverbal clause
3. One tense, aspect, modality and polarity value that is shared by all verbs. This is normally only marked on one verb, but can be marked on all of them
4. No markers of subordination, coordination or any kind of syntactic dependency
5. The verbs share at least one core argument
6. There is only one grammatical subject
7. The construction refers to a single event

However, exceptions have been noted for almost all these criteria. For instance, Ewe, which is otherwise a classic example of a serializing language, allows independent marking of aspect on verbs within SVCs (Ameka 2006a). Only monoclausality has been unquestioned, though it is clearly an insufficient criterion on its own. This list of criteria is, therefore, often taken as identifying an SVC prototype from which individual cases may deviate slightly (e.g. Foley 2010a). The characteristics of SVCs in Avatime and more generally cross-linguistically are discussed further in Chapter 3.

## 1.5 Serial verb constructions and events

As noted in the previous section, referring to a single event is often taken to be a defining criterion of SVCs. This criterion is both central to the definition of SVCs and highly problematic. Durie (1997) suggests it may be the only property shared by all SVCs and Bisang (2009) argues that it is the basic property from which all other prototypical properties of SVCs can be derived. At a syntactic level, reference to a single event links SVCs with single clauses, the lack of subordination and coordination, and the presence of a single grammatical subject. At a semantic level, it motivates the single values for verbal categories such as tense, aspect, modality, and polarity, as well as the need for shared arguments. The use of several verbs to

form a single predicate referring to a single event is thus the core of the SVC prototype. However, the single event criterion is also the most problematic aspect of SVC definitions. It has been criticized for its vagueness and the difficulty of evaluating it in a non-circular manner (e.g. Crowley 2002; Foley 2010a; Pawley 2011; Senft 2008; van Staden & Reesink 2008). Some researchers have also suggested the reverse relationship and claimed that SVCs in fact refer to multiple events (e.g. Baker & Harvey 2010).

Very few studies have explicitly set out to investigate the relationship between SVCs and conceptual events. Although many scholars have acknowledged the need for further work in this area, they have often dismissed it as too difficult (e.g. Foley 2010a; Pawley 2011; Senft 2008). In practice, most scholars describing SVCs in particular languages have relied on intuition, translation, and, less commonly, cultural restrictions on SVC use to argue that they refer to single events (Crowley 2002; van Staden & Reesink 2008). There have, however, been two previous dedicated investigations of the relationship (Baker & Harvey 2010; Givón 1991). These two studies used different methods and came to opposite conclusions.

Givón's (1991) approach was based on the idea that pauses in speech occur when speakers are encoding the next unit of speech (Goldman-Eisler 1968). This motivated his use of pauses as an indication of boundaries between conceptual event units, since people would encode different conceptual events separately. In a collection of narratives from three serializing languages of Papua New Guinea, he found pauses were no more likely to occur between verbs in SVCs than within single lexical items, and they were much less frequent than pauses between clauses. Givón thus concluded that SVCs describe single events. This was a pioneering study. It was the first (and for a long time, the only) to explicitly investigate the long-discussed relationship between SVCs and single events. It has, however, been criticised (e.g. Pawley 2011) for assuming connections between linguistic and conceptual structure, which although intuitive have not been demonstrated. Namely, it assumes individual lexical verbs describe single events and that pauses in speech are an accurate indicator of event boundaries.

Baker and Harvey (2010) were the second to explicitly investigate the relationship between single events and SVCs. They investigated the kinds of semantic structures and meanings expressed by SVCs and single verbs across a sample of languages. They found the meanings expressed using SVCs often differed from those expressed by single verbs and had more in common with meanings

expressed by complex clauses. For instance, causation of transitive actions and instrumentals were not generally expressed using single lexical verbs but were commonly expressed with SVCs and complex clauses. They concluded that many SVCs do not describe single events. While it takes a very different approach and comes to the opposite conclusion as Givón, this argument runs into similar problems in that it also is based on the assumption that individual lexical verbs describe single events.

These two studies used different methods and came to opposite conclusions, despite the shared comparison to lexical verbs. Rather than resolving the question of whether SVCs refer to single events, they highlight the problems involved in investigating this question. The claim that SVCs refer to single events has long been criticized for its vagueness regarding what is meant by a single event. This criticism is clearly justified given that the two previous studies interpret the notion of a single event very differently without much discussion: Givón investigated event *tokens,* while Baker and Harvey investigated event *types*. The other common criticism concerns the difficulty in evaluating the claim. Both studies run into this problem due to the circularity of comparing SVCs with single verbs and the assumption that single verbs refer to single events. What we need is a language-independent measure to test the relationship between SVCs and single events and a clearer, more explicit account of what is meant by 'single event' and what it means for an SVC to refer to one. This is what this thesis aims to do.

## 1.6 Data and ethics

The linguistic description is based on a corpus of Avatime collected by Saskia van Putten and myself over a combined 22 months in the field from 2008 till 2013. This corpus contains conversations, interviews, narratives, procedural texts, and elicited stories such as pear and frog stories. There is a broad range of social contexts from informal family discussions through to formal community meetings. It also includes a broad range of speakers including women, men, and children, with an age range from 4 through to 97 years old. This corpus was supplemented with elicitation sessions, typically held with one of five main consultants.

The co-speech gesture study in Chapter 4 examines narratives from four elderly Avatime speakers. These people and their narratives are described further in Section 4.2.1. In contrast, teenage participants were chosen for the studies in Chapters 5 and

6. This was because the social structure of the Avatime communities meant it was very difficult to recruit sufficient numbers of adult participants for these studies. Many Avatime adults leave the village for work, at least during the day, thus leaving a much larger proportion of children and elderly in the community than middle-aged adults. While elderly participants were available, they were not as well-suited to these types of experiments as they often had problems with their eyesight. While there were several differences noted between the speech of the teenagers and the adult speakers, these differences were mostly phonological, morphological, or lexical (see Chapters 2 and 3 for further discussion) and did not appear to extend to the syntactic properties of interest in the studies. Thus their behaviour is not predicted to be markedly different to that of adult speakers in these respects. The possibility of differences between the experiments due to age is discussed again in the Conclusions in Chapter 7.

Permission was obtained from appropriate authorities in all communities as well as from the individual participants and participating schools. Permission was obtained from the Avatime chiefs. Approval was granted by the Australian NSW Department of Education and Communities to conduct research in NSW high schools (SERAP 2011174). Approval was also granted by the Radboud University Ethics Committee for carrying out the research in the Netherlands (ECG2013-1304-097).

Participants were compensated according to local norms and in consultation with appropriate authorities. Dutch participants were compensated monetarily following the standard rates of the MPI for Psycholinguistics. English participants were also compensated monetarily according to these same rates after consultation with teachers in Australia. Avatime experimental participants were initially compensated monetarily at a rate suggested by community leaders. However, following the subsequent request of community and school leaders I shifted to compensating participants with school supplies of equal value. Avatime speakers who contributed to the corpus were offered gifts, and more regular consultants were paid at an agreed hourly rate. Beyond these direct compensations to participants, I strove to give back to the communities in other ways. I gave talks to participating schools in Australia, the Netherlands, and Ghana. There was also a desire among members of the Avatime community to preserve their cultural knowledge and in particular the knowledge of older people. This led me to focus my recordings, wherever feasible, towards documenting this knowledge. For instance, this was the main motivation for the topics chosen for the monologues studied in Chapter 4.

All recordings (from the corpus, elicitation sessions, and experiments) as well as experimental stimuli are archived at The Language Archive, which is located at the Max Planck Institute for Psycholinguistics. This archived material and can be found at https://hdl.handle.net/1839/00-0000-0000-0016-AA18-E@view. The Avatime linguistic examples in this thesis typically include a reference to the recording they are taken from. However, examples occurring in multiple elicitation sessions are not generally given a specific reference. These references align with the system of file naming in the archive. The reference occurs at the end of the example between brackets. The reference consists of a genre or topic identifier and the date in YYMMDD format. In many cases it also includes the initials of the speaker(s) and a time reference in MM:SS. Some examples of genre and topic identifiers used are *elic* for elicitation, *conv* for conversation, *frog* for frog story, and *folkstory* for folk stories.

## 1.7 Structure of the thesis

This thesis investigates the relationships between event segmentations in language and thought via a detailed case study of serial verb constructions in Avatime. This focus on one language and one construction type allows the investigation to delve into the details of the relationship with a solid grounding in the linguistic properties. It also leaves room for a broad investigation across different aspects of conceptualization, utilizing both gesture analysis and behavioural experiments to investigate the relationship at different levels: thinking-for-speaking, perception, and memory. Focusing on a language such as Avatime also has the advantage of extending the study of event segmentation beyond speakers of English and other Literate, Official, with Lots of users (LOL) languages (Dahl 2015) or Western, Educated, Industrialised, Rich, and Democratic (WEIRD) populations (Henrich, Heine & Norenzayan 2010).

The thesis is broken into two main parts: language description (Chapters 2 and 3) and experimental investigation (Chapters 4, 5 and 6). Chapter 2 provides a sketch grammar of Avatime. This extends previous descriptions of the language and provides a basis for the rest of the thesis. Not only does it assist the reader to interpret the example sentences throughout the thesis, but it provides a description of the general language ecology in which the serial verb constructions fit. Chapter 3 provides a detailed description of Avatime serial verb constructions. It identifies

three subtypes of SVC in Avatime and contrasts Avatime serial verb construction properties with those in languages spoken nearby and further afield.

This linguistic description is followed by three experimental chapters, each focusing on a different aspect of conceptual event segmentation. Chapter 4 examines the co-speech gestures of four Avatime speakers in order to investigate the relationship between event units during thinking-for-speaking and the various multi-verbal constructions (including SVCs) in Avatime. Chapter 5 investigates the relationship between SVCs and event segmentation during perception. It asks whether familiar events, which may be more likely to be described with SVCs, might in fact be segmented more holistically or coarsely during perception. Conversely, it also investigates whether the use of SVCs may influence the way events are segmented. Chapter 6 investigates whether events described using SVCs are remembered as single events.

# 2 Avatime Grammar Sketch

## 2.1 Introduction

The purpose of this chapter is to provide background information on the Avatime language (Glottocode: avat1244, ISO 639-3: avn) and the people who speak it. Funke's (1909) short description is the only previous general description of the grammar and it is not sufficient for the purposes of the present thesis. Other previous work on the language has focused on describing particular aspects of the grammar such as phonology (Ford 1971a; Maddieson 1998; Schuh 1995a), noun classes (Ford 1971b; Schuh 1995b), aspect and mood (Defina 2009), syntax (Ford 1971a), and information structure (van Putten 2014a). The sketch grammar provided here is based on research conducted with Saskia van Putten (see van Putten 2014a) during a combined 22 months of fieldwork between 2008 and 2013. Its primary contributions are in its descriptions of: the tonal system, with three rather than four level tones (primarily based on van Putten's work); vowel sequences and the processes used to reduce them (primarily based on my own work); noun phrases (van Putten); verbal morphology, including infinitivization and nominalization (myself); and the syntax of simple and complex clauses (van Putten and myself). The discussions of vowel sequences (Section 2.2.4.3), verbs (Section 2.7), and simple (Section 2.10) and complex clauses (Section 2.11) are especially relevant for this study of serial verb constructions.

### 2.1.1 Avatime: the name

The name Avatime refers to both the language and the people who speak it. It comes from Ewe, the local regional language and is commonly used by non-Avatime people and in academic discussions of Avatime language and culture. The term is not considered derogatory and is commonly used by Avatime speakers when using other languages such as English or Ewe, for instance in the "Avatime roots" Facebook group.

In Avatime, the language is referred to as *sị̀yàsɛ* or *sìdemè(se)*. *Sị̀yà* is the Avatime word for language, with the addition of the definite suffix *-sɛ* it means 'the language' and is generally interpreted as Avatime. *Sìdemè(se)* is a new word which has developed during the course of the Ghana Institute for Linguistics, Literacy and Bible Translation (GILLBT) project on Avatime. The word appears to be based on the word for 'back' commonly used in reference to the people and the land, ostensibly due to their history as the people who stayed back during a migration. It

can be analysed as *sì-,* the noun class prefix used with languages, + *de,* the root for 'back', + *mè* 'inside', and optionally the definite suffix *–se*. In 2008, most speakers referred to their language as *sị̀yàsɛ* and many did not recognise the term *sìdemè(se)*. Since then, *sìdemè(se)* has become more widely accepted and is used about as frequently as *sị̀yàsɛ*.

The Avatime refer to themselves as *kedɔ́nɛ̀* (*kede* 'back' + *ɔ́nɔɛ̀* 'person') for an Avatime man, *kededze* (*kede* 'back' + *odze* 'wife') for an Avatime woman, and *kedánà* (*kede* 'back' + *bánɔà* 'people') for Avatime people. The area they live in is referred to as *kedeamè* 'in the back' (*kede* 'back' + *a* 'definite' + *mè* 'inside').

### 2.1.2 The people and their setting

The Avatime traditional area is located about 50 kilometres north of the regional capital Ho, close to the border with Togo. There are eight Avatime villages: Amedzofe, Biakpa, Dzogbefeme, Fume, Gbadzeme, Old Dzokpe, New Dzopke, and Vane (see map in Figure 2.1).

According to their oral history, the Avatime people migrated to their current location from Ahanta in the South-West of Ghana (see also Brydon 2008; Heine 1968). The migration ended in the lands of the Baya people. The incoming migrants subdued the Baya inhabitants and then mixed with the Baya survivors dispersing them throughout the Avatime villages (Brydon 2008). According to the oral histories, the Baya people already spoke the Avatime language and the immigrants adopted it. This is also corroborated by the form of their name which consists of the noun class prefix used for people *ba-* and the root *yà* which is the same root as in the endonym *sị̀yàsɛ*.

There are approximately 15,000 Avatime speakers. The analysis of the 2010 census was not yet complete during my final field trip. However, working from the 2000 census and assuming the Avatime population changed at a similar rate to the national population, there should have been approximately 10,000 people living in the Avatime villages as of 2010. Since the population has likely increased since 2010 and there are a substantial number of Avatime people living outside the traditional area, I estimate the total population to be around 15,000. This is significantly less than the estimation of 24,000 provided by Ethnologue (Lewis, Simons & Fennig 2014) which is based on a survey conducted by GILLBT in 2003.

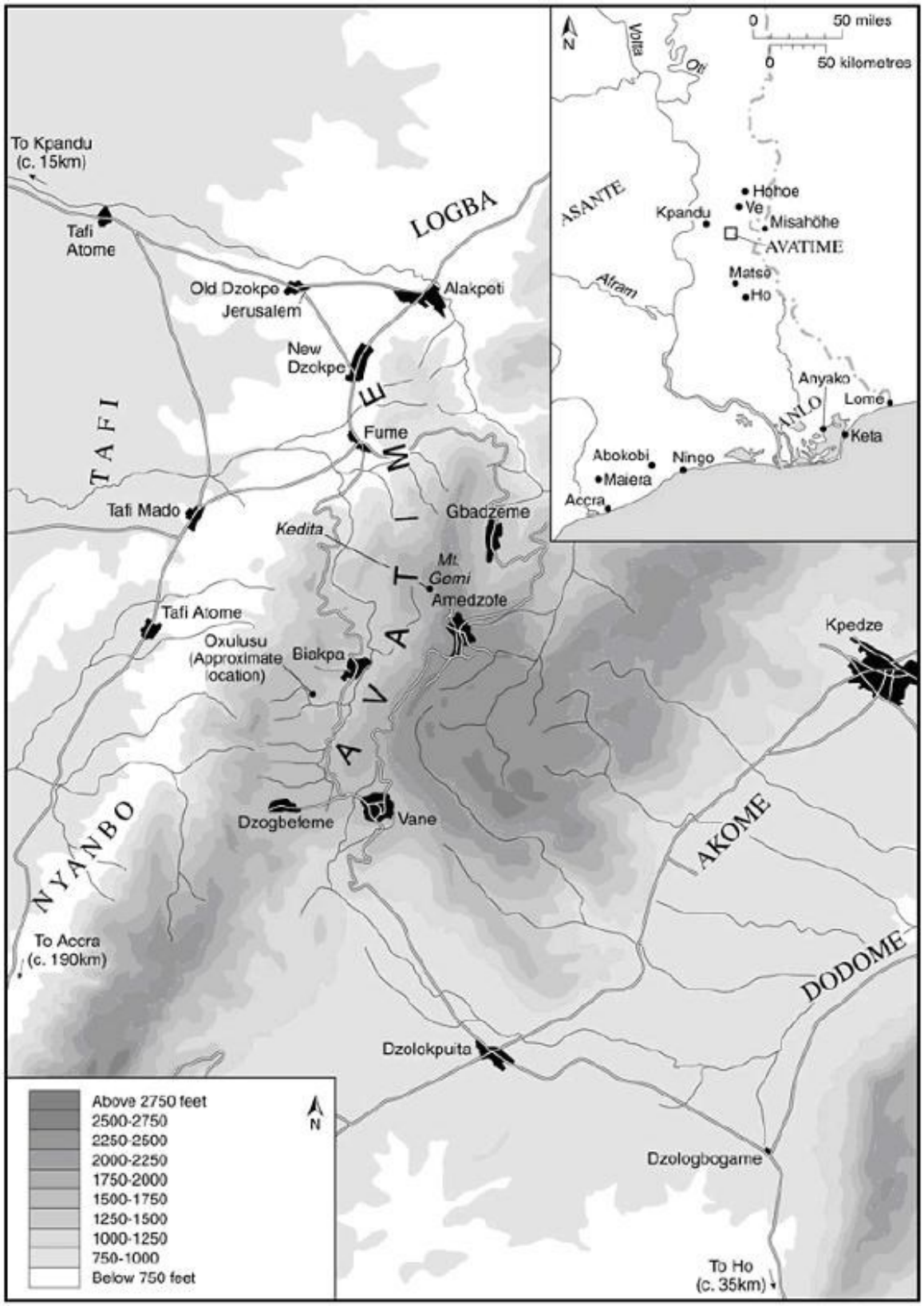



Figure 2.1: Map of the Avatime traditional area and some of the surrounding villages. The inlay shows the location of the traditional area in relation to the wider area (Source: Brydon 2008:25)

Most Avatime people are subsistence farmers, though there are generally a number in each family who take up other professions. Each person has several plots of land in the areas surrounding the villages. The main staple crops are cassava, maize, yam, and rice. Cocoa is also grown as a cash crop. The Avatime people are known for a tradition of rice cultivation and there were many cultural traditions tied to the farming of rice (Brydon 1981). In the last few decades these traditions have largely been lost, though in 2010 the Paramount Chief introduced annual rice festivals in an attempt to revive some of these cultural practices.

The Avatime people currently have a chieftaincy system modelled on the Ashanti chieftaincy system. It appears to have been adopted relatively recently, likely in the early 19th century (Brydon 1981; Brydon 2008). Each Avatime village has its own male and female chief (both referred to as *okusie*). The entire Avatime area is under the domain of a paramount male chief and a paramount female chief (both referred to as *osie*). The village of Vane is the seat of the paramountcy and the paramount chiefs reside there. More traditionally, each village is divided into clans (*akpɔ*, sg. *ḷikpɔ*) which in turn consist of smaller patrilineal groups (*iku*, sg. *oku*). The elder of each *oku* is in charge of land allocation for farming and housing within each group. For more information regarding Avatime social organization see Brydon (1981; 1996; 2008).

There is dialectal variation between the Avatime villages. The full extent of this variation and the number of distinct dialects has not been fully investigated. I have worked in four of the Avatime villages: Amedzofe, Biakpa, Gbadzeme and Vane, with most of my work conducted in Vane. During this time (2008-2013) I have noted phonological differences between the dialects, for instance intervocalic *w* in Amedzofe and Gbadzeme corresponds to β in Biakpa and is lost in Vane. I have also noticed morphological differences, for instance there is a negative marker in Amedzofe (*bí-*) which has been lost in Vane (Defina 2009), and lexical differences, for instance the word for pepper is commonly provided by speakers as an example of the differences between the dialects as it is said to be different in each village.

The Avatime community is surrounded by speakers of other languages. They are situated within the large Ewe speaking region of Ghana and Togo and Ewe is commonly used as a lingua franca. Ewe is also used in the Avatime community in church and for the first three years of school. There are also three smaller language groups bordering the Avatime lands: Logba, Nyagbo, and Tafi. Nyagbo and Tafi are

both closely related to Avatime within the Ka branch of the Ghana-Togo Mountain languages. Logba, a member of the Na branch of the Ghana-Togo Mountain languages, is more distantly related. A few Avatime people also speak one of these languages. The national language of Ghana is English and it is often used in the media (newspapers, radio and TV), in school, and when people travel outside of the Volta region. Most Avatime people under 40 have attended school and learnt English. Avatime remains the main language within the Avatime villages. It is used at home, on the street, at the local markets, in public meetings, and traditional ceremonies.

### 2.1.3 Linguistic classification

Avatime is one of the Ghana-Togo Mountain languages. These languages are generally believed to be part of the Kwa subgroup of the Niger-Congo language family, though they are also known for their typological differences from the surrounding Kwa languages (Blench 2009; Heine 2008; Stewart 1989; Williamson & Blench 2000). In earlier literature, they have been referred to as *Togorestsprachen* 'Togo Remnant Languages' (e.g. Heine 1968) and Central Togo Languages (e.g. Kropp Dakubu & Ford 1988).

The relations between the Ghana-Togo Mountain languages and the rest of the Kwa languages are not clear, mostly because the Kwa language family in general is not well-defined (Aboh & Essegbey 2010; Stewart 1989; Williamson & Blench 2000). Heine (1968) divided the Ghana-Togo Mountain languages into two sub-groups – Na and Ka – with largely flat substructures. The names for the groups are based on the reconstructed word for 'meat' in each sub-group. Avatime is in the Ka sub-branch along with Ahlo, Akebu (Kebu), Animere, Ikposo (Kposo), Nyagbo, Tafi, and Tuwuli (Bowili). This sub-division has generally been adopted in subsequent work with a few more details added to the subclassifications (e.g. Ameka 2009; Bobuafor 2013; Dingemanse 2011; Dorvlo 2008; Kropp Dakubu in press; Kropp Dakubu & Ford 1988; van Putten 2014b; Williamson & Blench 2000). However, Blench (2009) has recently suggested dividing the primary Na vs Ka distinction into four smaller clusters. There is more disagreement regarding how and where the Ghana-Togo Mountain languages connect to the rest of Kwa (e.g. Blench 2009; Heine 1968; Kropp Dakubu in press). However, the most recent evidence presented by Kropp Dakubu (in press) presents a convincing argument that they are connected as a single group with the Tano languages as their closest relatives. Under all analyses,

Avatime's closest relatives are Nyagbo and Tafi. Figure 2.2 shows the current best estimate of their family structure.

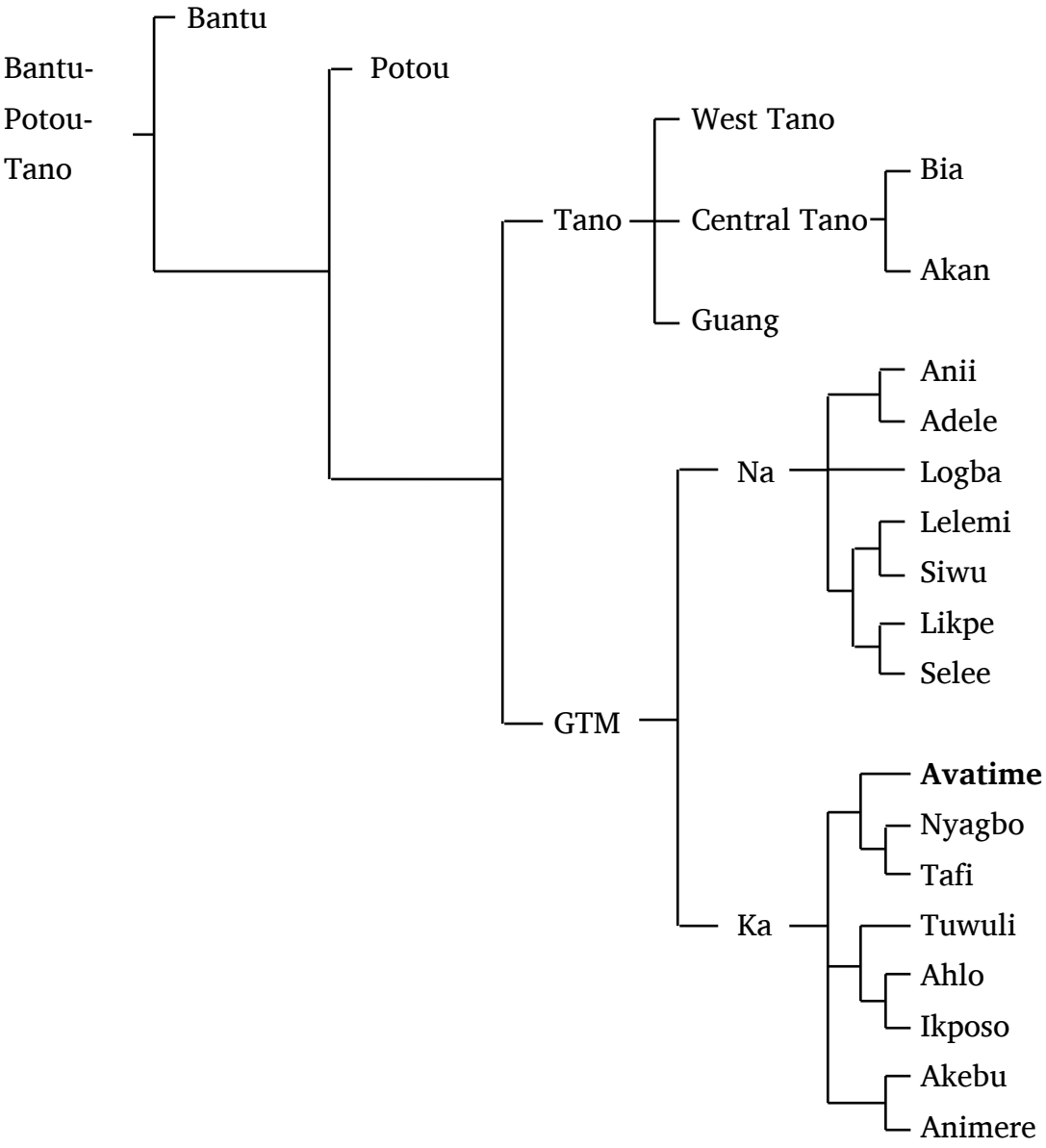



Figure 2.2: Current best estimate of the relations between the Ghana-Togo Mountain and other Kwa languages based on Blench (2009), Stewart (1989), and Kropp Dakubu (in press). Branch lengths are not indicative.

## 2.2 Phonology

### 2.2.1 Phoneme inventory and orthography

#### 2.2.1.1 Orthography

The orthography used in this thesis for Avatime is based on the Ewe orthography, since the phoneme inventories of Avatime and Ewe are very similar and Ewe is the dominant language in the region. The orthography used here is also very similar to the orthography recently developed for Avatime by the GILLBT team working on the language. The main differences between the two orthographies are as follows: 1) palatal affricates are written /ky/ and /gy/ in the GILLBT orthography as opposed to /tsy/ and /dzy/ in the present orthography; 2) retracted tongue root (–ATR) high vowels are not distinguished from their advanced tongue root (+ATR) counterparts in the GILLBT orthography; 3) tone is not regularly marked in the GILLBT orthography.

#### 2.2.1.2 Consonants

Avatime has 29 consonants, listed in Table 2.1. The consonants between brackets only occur in loan words from Ewe. Where the orthography deviates from IPA, the IPA symbol is provided between square brackets.

Table 2.1: Consonant chart

| | bilabial | labio-dental | alveolar | palatal | velar | labial-velar |
|---|---|---|---|---|---|---|
| stop voiceless | p | | t | | k | kp |
| stop voiced | b | | d (ɖ) | | g | gb |
| fricative voiceless | (ƒ [ɸ]) | f | s | | x | xw [xʷ] |
| fricative voiced | ʋ [β] | v | z | | h [ɣ] | hw [ɣʷ] |
| affricate voiceless | | | ts | tsy [ʧ] | | |
| affricate voiced | | | ʣ | dzy [ʤ] | | |
| nasal | m | | n | ny [ɲ] | ŋ | ŋw [ŋʷ] |
| oral sonorant | w | | l/r | y [j] | | |

Most previous research on Avatime makes a distinction between alveolar and palatal affricates (Ford 1971a; Funke 1909; Heine 1968; Kropp Dakubu & Ford 1988). Schuh (1995a), however, claimed that the place of articulation varied from speaker to speaker. My research has found a phonemic distinction but one which is being lost by younger speakers. Older speakers make a clear distinction with

minimal pairs such as *tsa* ‘to meet’ vs. *tsya* ‘to donate, dash’. Younger speakers, up to about 40 years old, do not make a distinction and often favour the palatal variant.

Ford (1971a) claimed that Avatime has labial-velar fricatives (xɸ and ɣβ). However, Maddieson (1998) convincingly showed these sounds are phonetically labialized velar fricatives ($x^w$ and $ɣ^w$), rather than truly doubly articulated fricatives. Phonemically, they fit within the labial-velar class of sounds which include the doubly articulated stops (kp and gb) and the labialized nasal ($ŋ^w$).

The alveolar sonorants *l* and *r* are in complimentary distribution. The *r* occurs only as the second consonant in clusters with an alveolar or palatal initial consonant, e.g. *trɛ* ‘go’ and *ɲyrɔ̀* ‘sink’. The *l* is used in all other environments: as a single onset and as the second consonant in clusters with initial labials, velars, and labial-velars, e.g. *ple* ‘descend’, *lị-vlɛ* ‘morning’, and *lị-kla* ‘stone’. Exceptions to this generalization occur in ideophones, e.g. *rị̀dị̀dị̀* ‘continuously’ and *prùdù* ‘fly’ and loan words e.g. *saprada* ‘onion’.

The voiced bilabial fricative /ʋ/ varies across the different dialects. In some dialects, such as the Fume and Amedzofe dialects, it is clearly phonemic. In other dialects, such as in Vane, it occurs as a variant of /b/ and /w/.

#### 2.2.1.3 Vowels

Avatime has nine vowels with no phonemic distinction in vowel length (see Table 2.2). Like many other West African languages, Avatime makes a distinction between vowels with advanced tongue root (+ATR) and vowels with retracted tongue root (–ATR) Following the Ewe orthography, ɛ and ɔ are used for the –ATR mid vowels. The –ATR high vowels are marked with dots underneath the vowel (ị and ụ).

Table 2.2: Vowel chart

| | front | | central | back | |
|---|---|---|---|---|---|
| | + ATR | –ATR | (–ATR) | + ATR | –ATR |
| high | i | ị | | u | ụ |
| mid | e | ɛ | | o | ɔ |
| low | | | a | | |

Previous work on Avatime has disagreed on the presence of –ATR high vowels. They were noted in early work, but not distinguished in the orthography (Funke

1909). They were then absent from Avatime descriptions for some time (Ford 1971a; Heine 1968; Kropp 1967). More recently, Maddieson (1998) showed there is an ATR distinction in Avatime high vowels and that they actively participate in the vowel harmony process (see Section 2.2.4.2). I have found the distinction is clearly maintained by older speakers of the language, with minimal pairs such as *kikù* 'yam' and *kɪ̣kʊ̀ʊ̣* 'item made of rubber or plastic'. The distinction does appear to be disappearing in the language though, and younger speakers – such as teenagers – will often pronounce –ATR high vowels as +ATR.

Vowels can be nasalized, though the nasalization is rare and only ever pronounced word-finally and sometimes not even then. Schuh (1995a) found only three nouns and a few verbs with unconditioned nasalization and concluded that while nasal vowels had previously existed in the language they were in the process of disappearing and, in fact, very near the end of this process. I have found nasalized examples of all the –ATR vowels: *tsyɪ̀̃* 'to tear', *sɛ̃* 'to leave', *lɪ̣klã* 'stone', *ɔ̀gɔ̀̃* 'coconut', and *ɪ̣sʊ̃ʊ̀* 'body'. In contrast, there are no clear examples of nasalized +ATR vowels. There are, however, some cases where *e* and *u* appear to have triggered nasalization of a consonant indicating that the vowel was previously nasalized. This evidence comes from the definite articles in class 2p and 3s (Section 2.5.3.1) which are pronounced with an /n/ when the root is nasal and /l/ when there is no nasal in the root (see also Section 2.2.4.1). This means that the vowels in *ì-tsre-nè* 'the okra' and *li-vu-nè* 'the nest' must be, or once have been, nasalized even though the nasalization is not audible even when the vowel occurs word-finally without the definite suffix. Since nasalization does not play a major role at present in Avatime and it is often hard to determine whether or not a vowel is nasalized, nasalization is only marked in phonetic transcriptions within this thesis.

### 2.2.2 Tone

Avatime has three contrastive level tones: extra-high (marked á), high (unmarked), and low (marked à). The low and high tones are the more frequent and occur on all word types and various affixes. The extra-high tone has a limited distribution. It doesn't occur on noun roots, with the exception of loan words, e.g. *ɔ́flága* 'flag', and roots based on ideophones, e.g. *ɔ́kʊ́ʊ́kʊ* 'chicken'. It occurs only rarely on verb roots and on some noun class prefixes depending on the root. There are some verbal prefixes which bear extra-high tone, frequently as the result of tone raising

processes (Section 2.2.4.4). Some minimal pairs distinguished only by tone can be seen in (1).

(1) a. *ní* ‘locative preposition’
*ni* ‘extinguish (fire)’
*nì* ‘and/with’
b. *sịyà* ‘hair’
*sìyà* ‘language’
c. *tsyi* ‘pour’
*tsyí* ‘turn’

Contour tones on a single vowel usually clearly result from floating tones or the merging of two syllables. There are two exceptions: the conjunction *lɛ̌* and the prefix *zɛ̌-* used in the habitual and recurrent aspects (Section 2.7.2). In fast speech, contour tones are often produced as flat high tones.

Previous work on Avatime has identified four level tones: low, mid, high, and extra-high (Ford 1971a; Schuh 1995a). All the words which were previously described with a mid tone were checked with six consultants. The words were tested in isolation as well as in sentences and where possible followed by low tones and high tones. None of the consultants produced a mid tone in any of these contexts. Therefore, the present analysis includes only three level tones.

### 2.2.3 Root and syllable structure

Roots are predominately monosyllabic. Noun roots with more than one syllable are predominately loan words. There are several disyllabic verb roots, though the second syllables of many of these roots appear to have been suffixes historically.

Syllables can have the following structures: CV, for example *tɔ* ‘cook’ and *lị́* ‘be.at’; V, word initially as in *obi* ‘child’, or as the second syllable in a root as in *tà.e* ‘a little’; CVN where the N is a nasal, this only occurs in particles and ideophones such as *kóŋ* ‘very’; and CCV where the second consonant must be an oral sonorant, for example *ŋwlịmìị* ‘write’ and *òtsre* ‘okra’.

It is not entirely clear whether Avatime should be analysed as having CVV type syllables. There are three potential CVV type syllables: 1) Some syllables are likely CVV underlyingly but realised as CCV (e.g. *ɔmwị* ‘goat’ may be *ɔmụị* underlyingly).

Schuh (1995a) suggested aspects of the vowel quality may be preserved in subphonemic differences in the glides produced, though further study is needed to verify this. 2) Some roots are pronounced as CVV in the Vane dialect (e.g. *ka-sɔ̀ị̀* 'basket' and *tàe* 'a little'). These roots can be better analysed as bisyllabic CV.V since they correspond to bisyllabic roots in other dialects, for instance in Amedzofe they are pronounced *ka-sɔwị̀* 'basket' and *tàve* 'a little'. It is likely that intervocalic *w* and *v* have recently been lost in Vane thus creating these unusual CVV structures. 3) Some prefixes have a CVV structure, such as the potential mood prefixes (e.g. *kíà-* 1p potential prefix) and some of the noun class prefixes on numerals (e.g. *tia-bà* 'two'). Some of these have clearly been derived from the fusion of two prefixes, such as the subject agreement prefix plus a prefix *a-* for the potential. As such they could be analysed as historically bisyllabic.

### 2.2.4 Phonological processes

#### 2.2.4.1 Nasalisation of consonants

Oral sonorants in suffixes become nasalised following a syllable with a nasal. For example *-lo/-lɔ* 'Class 2 singular definite' becomes *-no/-nɔ* in *ɔ̀-mɔɛ-nɔ̀* 'the orange' and ɔ̀-g̰ɔ̰̀-nɔ 'the coconut'.

#### 2.2.4.2 Vowel harmony

Avatime has a system of vowel harmony based on the advanced or retracted tongue root distinction (+ATR or –ATR) (see Section 2.2.1.3). Prefixes and suffixes are unspecified for ATR value and take on the value of the root. In most cases all vowels within a root have the same ATR value, though ideophones and loanwords generally do not require vowels to agree. Compounded roots also maintain their own ATR values. If a root contains vowels with different ATR values, prefixes will harmonise with the first and suffixes with the final vowel.

There are five vowel harmony pairs: /i, ị/, /u, ụ/, /o, ɔ/, /e, ɛ/, and /e, a/. The +ATR vowel /e/ has two possible –ATR variants. The variant used depends on the morpheme: in suffixes /e/ always alternates with /ɛ/ and /a/ is invariant; in some prefixes /e/ alternates with /ɛ/ such as the class 1 plural subject agreement marker *be-/bɛ-*; in other prefixes /e/ alternates with /a/ as in the class 1 singular subject agreement marker *e-/a-*. This quirk in the system is likely due to the historical loss of a +ATR low vowel which would have alternated with /a/.

#### 2.2.4.3 Vowel sequences

Schuh (1995a) noted that vowel sequences are not allowed in Avatime. This restriction is stricter in some dialects of Avatime than in others and seems least restrictive in the Vane dialect where the loss of intervocalic /w/ and /ʊ/ has generated several vowel sequences (see Section 2.2.3). Nevertheless, all dialects make use of the following processes to avoid vowel sequences to some extent. In each case, there are several possibilities and it is not possible to predict with certainty how a particular vowel sequence will be realised.

Across word boundaries, sequences of vowels are typically avoided by the insertion of a glottal stop or by deletion of one of the vowels, as in (2). Both strategies can be used across a broad range of syntactic linkages, for instance between a possessor and a possessed noun, verb and object, subject and verb, and between verbs in serial verb constructions and nonfinite subordinate constructions. Vowel deletion across clause boundaries does not appear to occur. If the first vowel is high or mid it may alternatively be converted into a glide, though this is rare and only observed in two cases both with a verb and its direct object.

(2) [ìselèlɛ]

| | |
|---|---|
| *ì-se = lè* | *ị̀-lɛ* |
| $C_2$p-tree = DEF | $C_2$p.PFV-be.at |

‘There are trees.’ (Expnew06_s2_130805_07:30)

Within words, the preferred strategies for avoiding vowel sequences are vowel deletion and glide formation. Glides can only be formed if the first vowel is high or mid, but otherwise there are no clear restrictions or tendencies regarding when each strategy will be used.

When a vowel is deleted, the remaining vowel can be, though isn’t always lengthened. The deleted vowel is generally the highest vowel in the sequence. This is not an absolute rule, however, and there are several cases where lower vowels are deleted. Schuh (1995a) reported that +ATR vowels are deleted in preference to –ATR vowels. I have not found this preference. Some morphemes appear to be privileged against deletion. In particular, the vowel prefix used in serial verb constructions (Chapter 3) appears to be highly resistant to deletion.

The above discussion holds for sequences of distinct vowels. When two vowels of the same quality are adjacent, they are consistently pronounced as a single long vowel. Because each vowel maintains its own tone, this is a major source of (apparent) contour tones (Section 2.2.2).

#### 2.2.4.4 Tone raising

Tone raising occurs in both verbs and nouns (Ford 1971a). The processes appear to be similar for nouns and verbs. Since raising in verbs is better understood this will be the focus of this section.

Low and high tones are either stable (if they do not undergo raising) or unstable (if they do). In tone raising contexts, if a verb with an unstable low or high tone is immediately followed by a high or extra-high tone, the unstable tone in the verb will raise and the subject agreement prefix tone will often also raise (3).

(3) a. Low tone followed by low tone (no raising)
*ma-ŋà* *blàli=e*
1s.PFV-eat plantain=DEF
'I ate plantain.'

b. Low tone followed by high tone (raising)
*má-ŋa* *kị-mịmị=ɛ̀*
1s.PFV-eat $C_4$s-cooked.rice=DEF
'I ate rice.' (Elic-RS_080902-2)

Whether or not tones in the verb root and prefix are raised depends on whether they are low or high and the aspect and mood. Unstable low tones in the verb and the tones in their prefixes are raised in the positive, but not in the negative perfective, and not at all in the progressive aspect. There is conflicting evidence in the potential and intentive moods: some consultants raise low tones and others do not. Verbs with unstable high tones fall into two classes. For some verbs, the tones are raised in the same contexts as unstable low tones and the tones in their prefixes are also raised, see example (4). For other verbs, only the tone of the verbs is raised, and it is raised in all aspects, moods and polarities, see example (5).

(4) a. High tone in the positive perfective (raising of verb and prefix)
*é-dzé* *li-nyi = nè*
$C_1$S.PFV-forget $C_3$S-name = DEF
'She forgot the name.'

b. High tone in the negative perfective (no raising)
*ó-dze* *li-nyi = nè*
$C_1$S.NEG.PFV-forget $C_3$S-name = DEF
'She did not forget the name.' (Elic-SIS-tone_100708_MiA)

(5) a. High tone followed by low tone (no raising)
*kị-tɔ* *blàli = e*
1p.PFV-cook plantain = DEF
'We cooked plantain.' (Elic-RS_080902-2)

b. High tone followed by high tone in the positive perfective (raising of verb only)
*kị-tɔ́* *kị-mịmị = ɛ̀*
1p.PFV-cook $C_4$S-cooked.rice = DEF
'We cooked rice.' (Elic-RS_080902-1)

c. High tone followed by high tone in the negative perfective (raising)
*ɔ́-tɔ́* *kị-mịmị = ɛ̀*
$C_1$S.NEG.PFV-cook $C_4$S-cooked.rice = DEF
'She did not cook rice.' (Elic-verbtone-100708-AB)

d. High tone followed by high tone in progressive (raising)
*ɛ̀ɛ́-tɔ́* *kị-mịmị = ɛ̀*
$C_1$S.PROG-cook $C_4$S-cooked.rice = DEF
'She is cooking rice.' (Elic-verbtone-100708-AB)

## 2.3 Nouns and noun classes

One of the distinguishing features of Ghana-Togo Mountain languages is their elaborate noun class systems. Nouns consist of a noun stem and a noun class prefix indicating gender and number. Avatime has seven genders, six consist of a singular and plural pair and one is used for mass nouns (see Schuh 1995b). Unlike many other Niger-Congo languages singular and plural pairings are regular in Avatime. Various numbering systems have been used for Avatime noun classes, (see Schuh

(1995b) for a review). Here Heine's (1968) system is used. The prefixes for all noun classes can be seen in Table 2.3.

Table 2.3: Noun class prefixes

| Noun class | Prefix | Examples | |
|---|---|---|---|
| 1s | o-/ɔ- | *o-bi* 'child' | *ɔ́-dzɛ* 'woman' |
| 1p | be-/ba- | *be-bi* 'children' | *bá-dzɛ* 'women' |
| 2s | ò-/ɔ̀- | *ò-gbe* 'rope' | *ò-gle* 'week' |
| 2p | ì-/ị̀- | *ì-gbe* 'ropes' | *ì-gle* 'weeks' |
| 3s | li-/lị- | *li-bi* 'seed' | *lị̀-glì* 'wall' |
| 3p | e-/a- | *e-bi* 'seeds' | *à-glì* 'walls' |
| 4s | ki-/kị- | *ki-bu* 'thorn' | *kì-bu* 'honey' |
| 4p | bi-/bị- | *bi-bu* 'thorns' | *bị̀-klì* 'toes' |
| 5s | ku-/kụ- | *ku-dè* 'road' | *ku-tse* 'death' |
| 5p | bè-/bà- | *bè-dè* 'roads' | *be-tse* 'deaths' |
| 6s | ke-/ka- | *ke-pe* 'house' | *ke-se* 'ground' |
| 6p | kù-/kù- | *kù-pe* 'houses' | *kù-bè* 'hunger' |
| 7 | si-/sị- | *sì-yà* 'language' | *si-se* 'soil, mud' |

The tone of most noun class prefixes varies depending on the root. Some prefixes always bear a low tone (marked in Table 2.3). The prefixes of class 1 singular and plural never bear a low tone: they are either high or extra-high. In general, extra-high tones on noun class prefixes are rare and have only been observed in class 1, 3 and 4, both singular and plural, and class 5 singular. Sometimes the tone on the prefix is the only way to distinguish between words (see the class 4 singular examples in Table 2.3). Three pairs of noun classes differ only in their tone: 1s & 2s, 1p & 5p, and 5s & 6p. The first two of these are always distinct, since 2s and 5p always bear a low tone while 1s and 1p never do. The final pair is not always distinct, since 5s can also bear a low tone[1].

[1] Another way to analyse the system is to say that all class 5 singular forms with a low tone and all class 6 plural forms actually constitute a single noun class and that these nouns sometimes pair with class 5 plural forms and sometimes with class 6

The noun classes have rough semantic groupings (see Schuh 1995b). Humans are clearly grouped together in class 1. Most borrowed nouns are also placed in class 1, though they often do not acquire a prefix. Verbal nouns are almost always in class 5. Liquid and mass nouns are in either class 7 or 5. Other patterns are not as clear, though there are some tendencies. For instance, most nouns relating to time are in class 3, and while animals are spread out over classes 1, 2, 3, and 6, those in class 6 tend to be wild.

Some roots occur in different noun classes with altered meanings. Examples can be seen in (6). Some clear regularities are that a shift to class 7 is used to generate mass nouns and class 6 is used as a diminutive. Some less productive regularities are the use of class 3 for fruits and plant parts and class 5 for abstract concepts. This raises the question, not tackled here, of whether roots have inherent gender or not.

(6)

| | |
|---|---|
| *o-bi* 'someone's child' (class 1) | *li-bi* 'seed' (class 3) |
| *kụ-lị* 'palm tree' (class 5) | *lị̀-lị* 'palm fruit' (class 3) |
| *ó-bu* 'bee' (class 1) | *kì-bu* 'honey' (class 4) |
| *ɔ-nṵ̀vɔ* 'child' (class 1) | *kụ-nṵ̀vɔ* 'youth' (class 5) |
| *lị-klamị* 'small stone' (class 3) | *sị-klamị* 'sand' (class 7) |
| *ɔ̀-wlà* 'arm' (class 2) | *ka-wlà* 'hand' (class 6) |
| *ò-se* 'tree' (class 2) | *ke-se* 'stick' (class 6) |

Some words occur in multiple genders and this can at times disturb the otherwise regular pairings of singular and plural classes. For instance, the words for 'knee' and 'ear' are acceptable in two different genders. The word for 'knee' is often in class 3, *lị̀-klamasị* (singular) and *a-klamasị* (plural). However, some speakers use class 4, *kị-klamasị* (singular) and *bị-klamasị* (plural). Speakers sometimes explain their choice by saying the knee is like a stone *lị̀-kla* or like an elbow *kị-ŋwlaku*. The word for 'ear' is either in class 5, *kụ-tụkpa* (singular) and *bà-tụkpa* (plural), or in class 3, (*lị-tụkpa* (singular) and *a-tụkpa* (plural). Here the difference seems to be whether speakers consider the class 3 forms to be restricted to animal ears or whether it can also be used for human ears. Deviations from the regular singular

---

singular forms. Since this analysis disturbs the otherwise clear singular and plural system in Avatime it isn't adopted in this thesis.

and plural pairings arise when a speaker uses the singular form from one gender and the plural form from the other gender.

Another way the regular singular/plural pattern can be disturbed is with mass nouns. A collection of objects can sometimes be viewed as a plural collection or as a mass, for instance a single feather is in class 5 singular (*kụ-sịsị̀*), but a collection of feathers is either in class 5 plural (*bà-sịsị̀*) or in the mass noun class 7 (*sị-sịsị̀*). There are also some nouns which are typically thought of as mass but occur with singular or plural prefixes and do not generally have a plural or singular counterpart, for instance *kì-bu* 'honey' in class 4 singular and *kù-bè* 'hunger' in class 6 plural.

## 2.4 Pronouns

### 2.4.1 Subject, Object, Possessive pronouns

Avatime has one main set of independent pronouns which is used for subjects, objects and possessors (see Table 2.4). Distinctions are made between first and second person and all of the noun classes in singular and plural. When referring to people in the third person the class 1 pronouns are used.

Table 2.4: Pronouns

| | singular | plural |
|---|---|---|
| first person | mɛ | blɔ |
| second person | wɔ | mlɔ |
| class 1 (third person) | yɛ | ba |
| class 2 | lɔ | lɛ |
| class 3 | lɛ | la |
| class 4 | kɛ | bɛ |
| class 5 | kɔ | ba |
| class 6 | ka | kɔ |
| class 7 | sɛ | |

When used in subject position or as possessive pronouns, pronouns occur in the forms cited in the table, all of which have –ATR vowels. This can be seen in (7) for a subject pronoun and (8) for a possessive pronoun.

(7) ***mlɔ*** *tsyɛ* *mlɛ-tá-tsì*
**2p.SBJ** ADD 2p.PFV-INT-grow
'**You** too, you will grow up.' (Folkstory_ET_110827-2)

(8) ***yɛ*** *ò-nugu*
**$C_1$s** $C_2$S-mouth
'**his** mouth' (Folkstory-crocodile_PKD_110924)

When pronouns are used for objects they cliticize to the verb and vowel harmony takes place. So if the verb ends in a +ATR vowel the pronoun will also be produced as +ATR, except in the case of *a,* which remains unaltered. For instance in (9a) the first person singular pronoun occurs as the object of a +ATR verb and is produced as +ATR and in (9b) the verb is –ATR and the pronoun is likewise produced as –ATR.

(9) a. *àfua* *é-te=me*
Afua $C_1$s.PFV-know=1s.OBJ
'Afua knows me.'
b. *àfua* *a-xwa=mɛ*
Afua $C_1$s.PFV-call=1s.OBJ
'Afua called me.' (Elic-tone2-names-enclitics_120904_SO)

### 2.4.2 Indefinite pronouns

Avatime also has a series of indefinite pronouns formed with the noun class prefix and the root *tɔ*, which is also used to form indefinite articles (Section 2.5.3.1). Example (10) shows both an indefinite pronoun *batɔ* 'some people' and an indefinite article modifying the object noun phrase.

(10) *lɛ̌* ***ba=tɔ*** *bɛ-na* *ɔ̀-kʊ́-tɔ=ɛ*
CONJ **$C_1$p.SBJ=INDEF** $C_1$p.PFV-reach $C_2$s-place-INDEF=CM
'And **some** have reached someplace.' (Avatime-history_110905_BB_11:11)

### 2.4.3 Contrastive

There are two contrastive pronouns: *mɔ* for the first person singular (11) and *yɔ* for the third person (class 1) singular (12) (see van Putten, 2014a). As can be seen in the examples they have a kind of 'as for me/him/her' type meaning which singles the referent out as distinct from the rest of the group.

(11) ***mɔ*** *má-tá-ba* *lo*
**1s.CTR** 1s.NEG.PFV-INT-come FP
'Me, I won't come.' (Chiefs-meeting_100619-03)

(12) ***yɔ,*** *ma-mɔ̀* *sì* *á-sɛ* *kóko* *tsyɛ* *de*
**$C_1$s.CTR** 1s.PFV-see COMP $C_1$s.PFV-leave already ADD FP
'As for her, I believe she has left already.' (Conv-rice_110411-3-3)

## 2.5 Nominal modifiers and the noun phrase

The word order of the Avatime noun phrase can be seen in (13). The noun is obligatory and all other elements are optional. An example with all of these slots filled can be seen in (14).

(13) noun – adjectives – numeral – determiner – intensifier

(14) *e-boe* *kpekpe* *tata=là* *kò*
$C_3$p-matter short $C_3$p.three=DEF only
'only the three short matters' (Folkstory_110409_AB-2)

### 2.5.1 Adjectives

Avatime adjectives are distinguished by their position, immediately following the noun or another adjective, and their lack of agreement morphology. Many adjectives are ideophones, for instance *plɔplɔ* 'soft', *wụdzyɛwụdzyɛ* 'bushy, as of hair or an overgrown farm', *pititi* 'white', and *tsyịmịtsyịmị* 'sweet' (see Section 2.8). The non-ideophonic adjectives consist of a small group of non-derived adjectives and a number of adjectives which have a reduplicated structure and thus were likely

derived from verbs. The non-derived adjectives I have encountered are: *bìdi* ‘big’ and *gba* ‘good’, and *sìsami*[2] ‘small’. An example can be seen in (15).

(15) *ɔ̀-nyɔ* *mɛ̀* *dzɛ∼dzɛ* *ò-nu* *kị-dɔ̀* ***gba***
$C_2$s-farm inside REDUP∼go $C_2$s.PFV-be $C_4$s-thing **good**
‘Farming is a good thing.’ (Conv-greenhouse_110408_SO-ViA)

Verbs can be turned into adjectives by reduplication. An example is shown in (16), where (a) shows the verb *kpa* ‘to dry’ and (b) shows the adjective *kpakpa* ‘dry’. Reduplication is also a part of the nominalisation process (see Section 2.7.6).

(16) a. *bị-tá-**kpa***
$C_4$p.PFV-INT-**dry**
‘They will dry.’

b. *a-wlàkpa* ***kpa∼kpa*** = *la*
$C_3$p-leaf **REDUP∼dry** = DEF
‘the dry leaves’

A number of adjectives have a reduplicated structure where there is no corresponding verb, for instance *vuvu* ‘new’ and *wɔwɔ* ‘green/fresh’. These adjectives may have been derived from verbs which were subsequently lost in the language.

Several adjectives have verbal counterparts that express similar meanings but are not related in form. Some examples are: *kemè* ‘be big’ and *bìdi* ‘big’, *pɛ̀* ‘be good’ and *gba* ‘good’, *lɔsị̀* ‘be black and *kpị̀kpị̀* ‘black’. If an adjective does not have a verbal counterpart, it can be used predicatively in a copula construction with the verb *lí* ‘be at’, as in (17).

[2] In fact, *sìsamị̀* looks like it is probably derived from a noun, since *si-* is one of the noun class prefixes and *-mi* is a suffix which can be used to form diminutives. However, it does not appear to be a noun any longer as it now must be nominalised via the same procedure as other adjectives, as in (18).

(17) *yɛ* *ke-plikpà* *ka-**lí*** ***sìsamị***
$C_1$s.POS $C_6$s-book $C_6$s.PFV-**be.at** **small**
'His book is small.' (Contrexp15_s4_120909)

Adjectives can be nominalised by adding a noun class prefix. This can be seen in (18), where the adjective *sìsami* 'small' is prefixed with the class 1 singular prefix resulting in the meaning 'the small one (candle)'. Some adjectives can be repeated to intensify them, as in (19).

(18) *ó-nyime* *e-ni* *cándel* *bìdi=yè*
$C_1$s-man $C_1$s.PFV-put.off candle big=DEF
*lɛ̌* *ɔ́-dzɛ* *a-sụ* ***ɔ-sìsamị=ɛ***
and $C_1$s-woman $C_1$s.PFV-light **$C_1$s-small=DEF**
'The man extinguished the big candle and the woman lit the small one.'
(Contrexp31_s4_120913)

(19) *ke-plikpà* ***bìdi*** ***bìdi*** *á-tɔ* *ká-tịnį́* *ɔ̀-kplɔ=nɔ̀* *abà*
$C_6$s-book **big** **big** $C_6$s-INDEF $C_6$s.PFV-be.on:LOC $C_2$s-table=DEF on
'A **very big** book is on the table.' (Contrexp19_s3_120909)

### 2.5.2 Numerals

The numerals one to six are monosyllabic and underived, see Table 2.5.

Table 2.5: Numerals

| | |
|---|---|
| -le | one |
| -bà | two |
| -ta | three |
| -ne | four |
| -tsu | five |
| -glò | six |
| -glòele | seven |
| ɔ̀gɔtV(V)bà | eight |
| ɔ̀gɔtV(V)le | nine |
| liɔfɔ̀ | ten |

The numerals seven to nine are compositional. The word *–gloele* ‘seven’ consists of the root *–glo* ‘six’ and the root *–le* ‘one’. The words for eight and nine consist of the verb *gɔ* ‘remain’ inflected for class 2 singular subject agreement followed by tV(V)bà ‘two’ for eight and tV(V)le ‘one’ for nine. The numerals eight and nine can thus be literally translated as ‘two remain’ and ‘one remains’.

The numbers one to seven have a prefix o-/ɔ- when the numeral is used in isolation for counting. When used within a noun phrase, they have a prefix tV(V)- where the vowel(s) are determined by the noun class of the head noun (see Table 2.6). The tone depends on that of the head noun’s noun class prefix: if the prefix on the noun is low, then the prefix on the numeral is low; otherwise it is high. The numerals eight and nine show this agreement prefix in their third syllable. The numeral ten does not show noun class agreement.

Table 2.6: Noun class agreement prefixes for numerals

| | singular[3] | plural |
|---|---|---|
| class 1 | to- | tie-/tịa- |
| class 2 | to- | ti-/tị- |
| class 3 | ti- | te-/ta- |
| class 4 | ti- | tu(i)-/tụ(ị)-[4] |
| class 5 | tu- | tie-/tịa- |
| class 6 | ti- | tu-/tụ-[6] |
| class 7 | ti- | |

The numerals twenty to ninety consist of the form *avì* followed by the numerals two to nine inflected for class 3 plural, e.g. *avìtabà* ‘twenty’, and *avìtegloele* ‘seventy’. Hundred is *alafa*, borrowed from Arabic *alf* ‘thousand’ probably via Ewe. Thousand is *akpe*, borrowed from Ewe.

[3] The singular forms only ever occur with *–le* ‘one’ so there is no ATR harmony alternation.

[4] The vowel for the class 4 and class 6 plurals sometimes changes to [i] when the numeral has [u] in the root, e.g. *titsu* rather than *tutsu* ‘five’.

To form composite numerals, the noun phrase coordinator *(a)nì* ‘and/with’ is used, as in *avìtene nì ṭiabà* ‘forty two’. An example of a more complex numeral in use can be seen in (20).

(20) *ki-nu* *cedi* *akpe* *alafa* *ṭia-ta* *nì* *avìteglò*
$C_4$S.PFV-be cedi thousand hundred $C_1$P-three and sixty
‘It is three hundred and sixty thousand (360,000) cedis.’ (Tribunal_100513-4)

### 2.5.3 Determiners

#### 2.5.3.1 Articles

Avatime has a definite and an indefinite article. The definite article is widely used and is often added to nouns in isolation. The indefinite has a specific indefinite interpretation much like English ‘some’ or ‘a certain’. The bare noun is used for a non-specific indefinite interpretation. Both definite and indefinite articles follow the noun and any adjectives or numerals.

Table 2.7: Definite and indefinite articles

| | definite | | indefinite | |
|---|---|---|---|---|
| | singular | plural | singular | plural |
| class 1 | -(y)e/-(y)ɛ | -a | ɔ-tɔ | a-tɔ |
| class 2 | -lo/-lɔ | -le/-lɛ | ɔ-tɔ | ị/ɛ-tɔ |
| class 3 | -le/-lɛ | -la | ɛ-tɔ | a-tɔ |
| class 4 | -(y)e/-(y)ɛ | -e/-ɛ | ị/ɛ-tɔ | ị/ɛ-tɔ |
| class 5 | -o/-ɔ | -a | ɔ-tɔ | a-tɔ |
| class 6 | -a | -o/-ɔ | a-tɔ | ɔ-tɔ |
| class 7 | -se/-sɛ | | ɛ-tɔ | |

The definite article is a single syllable which can have either a CV or V structure and depends on the class of the head noun (see Table 2.7). It bears polar tone: so it is high if the preceding syllable is low, and low if the preceding syllable is high or extra-high. It is cliticized to the preceding word and the vowel harmonises with the ATR value of the preceding syllable. When the definite article consists of a single vowel, it creates a vowel sequence with the preceding vowel. This sequence is either maintained or, more typically, reduced via glide formation or vowel deletion

(Section 2.2.4.3), though the definite article vowel is privileged and unlikely to be deleted. Example (21) shows examples of nouns with and without the definite article.

| | | | |
|---|---|---|---|
| (21) | a. | [libì] | [libìle] |
| | | *li-bì* | *li-bì=le* |
| | | $C_3$s-wound | $C_3$s-wound=DEF |
| | | ‘a wound’ | ‘the wound’ |
| | b. | [kụlị] | [kụlị̀ɔ̀] |
| | | *kụ-lị* | *kụ-lị=ɔ* |
| | | $C_5$s-palm.tree | $C_5$s-palm.tree=DEF |
| | | ‘a palm tree’ | ‘the palm tree’ |
| | c. | [balị] | [balaà] |
| | | *ba-lị* | *ba-lị=a* |
| | | $C_5$p-palm.tree | $C_5$p-palm.tree=DEF |
| | | ‘palm trees’ | ‘the palm trees’ |

The indefinite article has the root *tɔ* with a prefix which agrees with the noun class of the head noun (see Table 2.7). This always leads to a vowel sequence, which in this case is typically resolved by the deletion of the previous vowel. The tone on the prefix depends on the tone of the noun class prefix on the head noun: if the noun class prefix is low, then the indefinite article prefix is also low; otherwise it is extra-high, as in (22).

| | | | |
|---|---|---|---|
| (22) | *lị-kla* + *ɛ́-tɔ* | [lịklɛɛ́tɔ] | ‘a certain stone’ |
| | *ɔ́-dzɛ* + *ɔ́-tɔ* | [ɔ́dzɔɔ́tɔ] | ‘a certain woman’ |
| | *kù-dze* + *ɔ̀-tɔ* | [kùdzeètɔ] | ‘a certain rat’ |

#### 2.5.3.2 Demonstratives

Demonstratives follow the noun and any adjectives and numerals. There is a proximal demonstrative *yè/yà* and a distal demonstrative *lɔ̀*. The proximal demonstrative takes the form *yè* when used with class 1 singular nouns and *yà* for all other noun classes. They agree with the head noun using the prefixes in Table 2.8. The prefix always bears an extra-high tone and this is also spread to the

preceding syllable. Examples can be seen in (23) and (24). Demonstratives can also be used as nominals, such as *kɔ́yà* 'this one' in example (25).

Table 2.8: Noun class agreement prefixes for demonstratives

| | singular | plural |
|---|---|---|
| class 1 | lí-/lɛ́- | bá- |
| class 2 | lɔ́- | lɛ́- |
| class 3 | lɛ́- | lá- |
| class 4 | kɛ́- | bɛ́- |
| class 5 | kɔ́- | bá- |
| class 6 | ká- | kɔ́- |
| class 7 | sɛ́- | |

(23) *trɛ ní* ***ke-pé*** ***kɛ́-ya*** *mɛ̀*
go LOC **$C_6$S-house** **$C_6$S-PROX** inside
'Go to **this house**.' (Folkstory_110409_AB-1)

(24) *mà-pɛ* *sị̀* ***li-bó*** ***lɛ́-lɔ̀*** *kɔ = ɛ*
1S.PFV-want COMP **$C_3$S-matter** **$C_3$S-DIST** CTR = CM
*lí-kí-dó* *ɔ-nɛnɛ* *kụ̀-sụ̀sụ* *mɛ̀*
$C_6$S.NEG.PFV-PROH-move.out:LOC $C_1$S-anybody $C_5$S-thought inside
'I want **that matter** not to leave anybody's mind.' (Chiefs-meeting_100619-3)

(25) ***kɔ́-yà*** *kɔ* *aní* *akpɛtɛshi* *lo*
**$C_5$S-PROX** CTR NEG akpeteshi FP
'As for this one (drink), it is not akpeteshi (a kind of liquor).'
(Conv-rice_110411-3-3)

### 2.5.4 Intensifiers

There are several particles which can occur at the end of noun phrases. These kind of particles have been called intensifiers in the previous literature (e.g. Ameka 2006b). They have functions such as quantifying, restricting, and contrasting. Some of them can also function as focus particles or as clause final particles. Some examples are *bóŋ* 'rather', *kɔ̀* 'only, just', *petee* 'all', *tsyɛ* 'also', and *tututu* 'exactly'

(see also van Putten 2014a). An example of the particle *tsyɛ* 'also' can be seen in (26). Multiple particles can occur together at the end of a noun phrase, as in (27).

(26) *ɔ-nụ̀vɔ̀ = ɛ* *sị̀sàmi = ɔ́-tɔ* ***tsyɛ*** *ɔ-lí́*
$C_1$s-child = DEF small = $C_1$s-INDEF **also** $C_1$s.PFV-be.at:LOC
*ke-pe = a* *mɛ̀*
$C_6$s-house = DEF inside
'A certain small child was **also** inside the house.'
(Famprob_110401_MeD-BeK_story)

(27) *lɛ̌* *mɛ* *sị̀* *lị̀-vlɛ = lɛ̀* ***tete*** ***bòŋwí***
and 1s.SBJ say $C_3$s-morning = DEF **only** **rather:FOC**
*tɔ* *mí́-panì = yɛ*
PURP $C_1$s.SBJV-greet = $C_1$s.OBJ
'And I said I would **rather** go and greet her **only** in the morning.'
(Conv-funeral_100528-7-1)

### 2.5.5 Possession

Possession is indicated by juxtaposition with possessor-possessum word order. The possessor can be a noun phrase (28) or a pronoun (29).

(28) *ɔ-gá* *lɛ́-lɔ* *a-sị̀a = nà*
$C_1$s-animal $C_1$s-DIST $C_3$p-horn = DEF
'the horns of that animal' (Frog_100719_DQ-PhA)

(29) *yɛ* *ò-nugu*
$C_1$s.POS $C_2$s-mouth
'his mouth' (Folkstory-crocodile_110924_PKD)

In the case of possessed kinship terms, the possessor pronoun is fused with the noun class prefix of the possessum, as in (30). Note that such a fusion can also occur in some other cases through the regular processes relating to vowel sequences, such as (29) which could be pronounced *yɛnugu*. However, in these cases the fusion is optional and either of the two vowels in the sequence may be elided. In the case of

inalienable kin possession the fusion is obligatory, regular, and occurs even when there is no vowel sequence, as in (30b). The fused possessor pronoun for inalienable kin possession is also present when the possessor is realised by a full noun phrase, as in (31).

(30) a. *yene*
*yɛ:o-ne*
$C_1$s.POS:$C_1$s-mother
'his mother'
b. *blaka*
*blɔ:ba-ka*
1p.POS:$C_1$p-father
'our fathers'

(31) *awu* ***ye**-ne*
Awu **$C_1$s.POS:$C_1$s**-mother
'Awu's mother' (Conv-funeral_100528-7-1)

The possessum can be pronominalized using the root *nɛ*, with the noun class prefix agreeing with the class of the possessum, as in the copularless sentence in (32).

(32) *kị-bɔ́* *kɛ́-yà* *wɔ* ***kị-nɛ***
$C_4$s-money $C_4$s-PROX 2s.POS **$C_4$s-POSM**
'This money is yours. (Elic-possessives_120906_PKD)

### 2.5.6 Conjoined noun phrases

Noun phrases can be conjoined using the conjunction/preposition *(a)nì* 'and/with'. An example of conjoined noun phrases can be seen in (33).

(33) *má-mɔ* *ɔ́-dzɛ* ***anì*** *ó-nyime* *ní* *ɔ̀-kplɔ=nɔ̀* *sụ*
1s.PFV-see $C_1$s-woman **and** $C_1$s-man LOC $C_2$s-table=DEF side
'I saw the woman and the man by the table. (Contrexp26_s3_120912)

Examples (34) and (35) show *(a)nì* used as a preposition: in (34) it is used to add an instrument and in (35) it is used to add an accompanier.

(34) *á-ŋa* *bị-dɔ̀mɛ* ***anì*** *gàtsyie*
$C_1$S.PFV-eat $C_4$p-thing **with** spoon
'She ate something with a spoon.' (Elic-QUIS-foc_100714_SO)

(35) *lɛ̌* *ɔ̀-lagɔ̀=lɔ=ɛ* *kị́à-klanị̀*
and $C_2$s-evening=DEF=CM 1p.POT-move.around
*ì-ma=nɛ* *mɛ̀* ***nì*** *blɔ* *bosòmi=a*
$C_2$p-town=DEF inside **with** 1p.POS candle=DEF.P
'And in the evening, we will walk around the town with our candles.'
(Chiefs-meeting_100619-03)

## 2.6 Adpositions and locative phrases

Locative phrases consist of a general locative preposition *ní* followed by a noun phrase or postpositional phrase. They can be used as adjuncts indicating the location – either spatial or temporal – of an event, or as oblique arguments of verbs of location, motion, and placement (for instance the verbs of putting and taking discussed in Chapter 6). The preposition *ní* is general and used with all locative phrases. The noun phrase indicates the reference object. The postposition indicates the search domain, i.e. the relevant part (inside, upper surface, etc.) of the reference object (for a similar account of locative phrases in Ewe see Ameka 1995).

The preposition *ní* is often reduced in casual speech, leaving the tone attached to the preceding syllable. Example (36) shows a locative phrase with overt *ní* (a) and with the *ní* reduced (b).

(36) a. *a-**trɛ*** ***ní*** *ke-pe=a* *mɛ̀*
$C_1$S.PFV-**go** **LOC** $C_6$s-house=DEF inside
'He went home.'
b. *a-**trɛ́*** *ke-pe=a* *mɛ̀*
$C_1$S.PFV-**go:LOC** $C_6$s-house=DEF inside
'He went home.' (Elic-S_081117_WO-1)

Avatime has five postpositions: *abà* ‘on’, *ese* ‘under’, *mè/mɛ̀* ‘inside’, *nu* ‘at the opening’, and *sụ* ‘near/next to’. Most of these are derived from nouns. The postpositions *abà* ‘on’, *ese* ‘under’, and *sụ* ‘near/next to’ are likely derived from the nouns *kabà* ‘top’, *kese* ‘ground’, and *ịsụ* ‘body’. The postposition *nu* ‘at the opening’ could be derived from the noun *ònugu* ‘mouth’ or it could also have been borrowed directly from Ewe. The postposition *mè/mɛ̀* ‘inside’ is the only one which does not appear to be derived from a noun and was likely borrowed from Ewe. This postposition is also unusual as it is preceded by a floating high tone and cliticizes to the preceding word where it participates in vowel harmony.

Nouns can also be used in place of postpositions to restrict the search domain. For instance, the nouns *ɔ̀tɔ̀nɔ* ‘front’, *kapɛ* ‘side’, and *kede* ‘back’ are often used for this function. Example (37) shows the noun *kede* ‘back’ used to narrow the search domain to the area behind a car.

(37) *xé gì ma-dɔ́ ò-hu=lò* ***ke-de=à***
when comp 1s.PFV-move.from:LOC $C_2$s-car=DEF **$C_6$s-back=DEF**
*mi-tsyí ple*
1s.SBJV-turn descend
‘When I come from behind the car, I should turn down.’ (Lego_081022_KA-RE)

A postposition is not required if the noun phrase already sufficiently specifies the location. This is especially common with place names, as in example (38).

(38) *mɛ̀ɛ́-trɛ* ***ní*** ***ò-ho=lò***
1s.PROG-go **LOC** **$C_2$s-Ho=DEF**
‘I am going to Ho.’ (Elic_120928_SO)

Deictic reference to a location is done using the words *níyà* ‘here’ and *nílɔ̀* ‘there’ formed by combining the preposition *ní* with the roots *yà* and *lɔ̀* also used for the proximal and distal demonstratives respectively (Section 2.5.3.2). There are several alternates for *nílɔ̀* ‘there’ (*níkɔlɔ̀*, *níklɔ̀, and níwlɔ̀*) suggesting it could also be derived from *ní* ‘LOC’ + *ɔ̀kɔ* ‘place’ + *lɔ̀* ’distal’. Even though these forms already contain the preposition *ní*, they are most commonly preceded by the preposition when they are used as oblique arguments or adjuncts after the verb, as in example (39). This suggests speakers do not perceive the *ní* as compositional here.

(39) *bíà-kɔ̀ nya **ní** **níyá** te*
$C_1$p.POT-take tie LOC **here** like.this
'They will tie it here like this.' (Illness_100616_SO-DS)

Their use as objects of the preposition *ní* suggests that *níyà* 'here' and *nílɔ̀* 'there' are nouns rather than adverbs. There are several other pieces of evidence supporting this. For instance, they can be used as subjects (40) and objects (41) of verbs. They can also head relative clauses (41).

(40) ***nílɔ̀*** ***ɛ-dra***
**there** $C_2$s.PFV-be.clean
'That place is clean.' (Fix-iron_110925_PKD)

(41) *lɛ̌ kíà-kɔ lùlɔ̀ **nílɔ̀** gì kí-ŋa li-wè=le*
and 1p.POT-take clean **there** REL 1p.PFV-eat $C_3$s-day=DEF
'And we will clean there where we celebrated the festival.'
(Chiefs-meeting_100619-03)

## 2.7 Verbs

Finite verbs in Avatime are obligatorily marked with a subject agreement prefix and one of a contrastive set of aspect and mood categories: perfective, progressive, habitual, potential, subjunctive, and imperative. They can also be marked by two optional aspect/mood categories: the recurrent and the intentive. There is no grammatical tense (Defina in press). It is also possible to mark the verb with a directional prefix (itive or ventive) and a comitative suffix. The structure of the full Avatime verb complex is shown in (42). Example (43) shows a verb with many of the slots filled in.

(42) Subject Agreement - (Negative) - Aspect/Mood - (Intentive) - (Recurrent) - (Prohibitive) - (Directional) - Root - (Comitative)

(43) *mɔ́-tá-zɛ̌-zɛ̌-panì=wɔ*
1s.NEG.PFV-INT-REC-IT-talk=2s.OBJ
'I will not be going to talk with you.' (Elic_081129_AB)

Positional and copular verbs form a distinct class in Avatime, signalled by the use of different subject agreement forms (Section 2.7.1). This class consists of the locative copula *lį́* 'be at', the equative copula *nu* 'be', and the positional verbs *dí* 'sit', *kpàsị* 'be in', *sụnụ* 'hang', and *tịnị* 'be on'. These verbs can only be used in the perfective aspect. In all other aspects and moods they are replaced by the suppletive copular verb *zè*. An example of this with the verb *kpàsị* 'be in' in the perfective and the progressive version with *zè* 'be' is shown in (44). The locative copula *lį́* 'be at' is further restricted in that it can only be used for present states, see example (45). It is the only verb in Avatime with sensitivity for temporal reference.

(44) a. *ɔ-nụ̀vɔ̀ = ɛ* ***ɔ-kpàsị*** *ní* *kà-sɔ̀ị = a* *mɛ̀*
$C_1$s-child = DEF **$C_1$s.PFV-be.in** LOC $C_6$s-basket = DEF inside
'The child is in the basket.'

b. *ɔ-nụ̀vɔ̀ = ɛ* ***èé-zè*** *ní* *kà-sɔ̀ị = a* *mɛ̀*
$C_1$s-child = DEF **$C_1$s.PROG-be** LOC $C_6$s-basket = DEF inside
'The child is still in the basket.' (Elic-R_081127_MM-1)

(45) a. ***ɔ-lį́*** *ní* *ɔ̀-va = nɔ̀*
**$C_1$s.PFV-be.at** LOC $C_2$s-Vane = DEF
'He is in Vane.'

b. ***e-zè*** *ní* *ɔ̀-va = nɔ̀*
**$C_1$s.PFV-be** LOC $C_2$s-Vane = DEF
'He was in Vane.' (Elic-R_081117_SO-1)

### 2.7.1 Subject agreement

There are three sets of subject agreement markers (Table 2.9). Each set contains a form for each noun class and person/number combination. The set used depends on the aspect, mood, and polarity value and in some cases on the particular verb. Set 1 is used in the positive perfective for most verbs. Set 2 is used with negatives, for the subjunctives of some verbs, and in the perfective with positional and copula verbs. Set 3 is used for the habitual and for the subjunctive of some verbs. The subject agreement prefixes are obligatory, except in the imperative which is signalled by their absence. If there is a lexical subject, the prefix serves an agreement function. If there is no lexical subject, the prefix has a pronominal like function.

Table 2.9: Subject agreement prefixes

| | Set 1 | Set 2 | Set 3 |
|---|---|---|---|
| 1s | me-/ma- | mo-/mɔ- | mí-/mḭ- |
| 1p | ki-/kị- | ku-/kụ- | kí-/kḭ |
| 2s | wo-/wɔ- | wo-/wɔ- | wú-/wṵ- |
| 2p | mle-/mlɛ- | mla- | mlí-/mlḭ- |
| $C_1$s | e-/a- | o-/ɔ- | i-/ị- |
| $C_1$p | be-/bɛ- | ba- | bi-/bị- |
| $C_2$s | è-/ɛ̀- | ò-/ɔ̀- | ì-/ị̀- |
| $C_2$p | ì-/ị̀- | ì-/ị̀- | ì-/ị̀- |
| $C_3$s | li-/lị- | li-/lị- | li-/lị- |
| $C_3$p | e-/ɛ- | a- | i-/ị- |
| $C_4$s | ki-/kị- | ki-/kị- | ki-/kị- |
| $C_4$p | bi-/bị- | bi-/bị- | bi-/bị- |
| $C_5$s | ki-/kị- | kù-/kṳ̀- | ki-/kị- |
| $C_5$p | be-/bɛ- | bà- | bi-/bị- |
| $C_6$s | ke-/kɛ- | ka- | ki-/kị- |
| $C_6$p | ki-/kị- | kù-/kṳ̀- | ki-/kị- |
| $C_7$ | si-/sị- | si-/sị- | si-/sị- |

In Set 1, the prefixes for first and second person singular and class 5 singular and plural and class 6 plural have variable tone. The tone is high if the verb has a low tone, and low otherwise, see (46). For the noun classes, the prefix will only be low if the prefix on the noun is also low.

(46) a. ***mè****-dzi* *dɔ̀mɛ*
**1s.PFV**-buy thing
'I bought something.'

b. ***me****-dò*
**1s.PFV**-move.out
'I went out.'

There are some patterns in the subject agreement forms indicating they were created by regular morphological and or phonological processes. At present,

however, the number of exceptions makes it difficult to determine exactly what the original basic forms were.

The class 1 singular subject prefix of a complement clause is marked as logophoric if it refers to the same entity as the subject of the matrix clause. This is done by adding *y* to the prefix. An example can be seen in (47) with a logophoric prefix from Set 3. There is no independent logophoric pronoun.

(47) *pɔ̀* *yɔ* *kɔ* *ɔ́-pɛ* *sị̀* ***yi-pe***
but $C_1$S.CTR CTR $C_1$S.NEG.PFV-want COMP **$C_1$S.LOG**.SBJV-tire
'But as for her, she doesn't want to get tired.'
(Conv-greenhouse_110408_SO-ViA)

### 2.7.2 Aspect and Mood

#### 2.7.2.1 Contrastive Aspect and Mood Categories

There are six contrastive aspect and mood categories: perfective, progressive, habitual, potential, subjunctive, and imperative. Only one of these can be used to modify a verb at a time. The contrastive aspect and mood markers are all combined to some extent with the subject agreement markers (see also Defina 2009).

##### 2.7.2.1.1 Perfective

The perfective is the most common aspect in Avatime (Defina 2009). It is indicated by using subject agreement prefixes from Set 1, or Set 2 if it is a positional or copular verb. The verb is otherwise unmarked. The perfective is used to refer to completed actions (48), states (49), and abstract situations such as in procedural descriptions (50).

(48) *me-dzì* *trɛ* *mɛ* *ke-pe=a* *mɛ̀*
1s.PFV-return go 1s.POS $C_6$s-house=DEF inside
'I've gone back to my house.' (Lego_081114_AB&WO_2)

(49) *ò-gbé* *lɔ́-yà* *è-dzè*
$C_2$s-rope $C_2$s-PROX $C_2$s.PFV-be.long
'This rope is long.' (Elic-S_0810271_SO)

(50) *kɔ* *mlɛ-trɛ* *ke-pe=a* *mɛ̀=ɛ*
then 2p.PFV-go $C_6$s-house=DEF inside=CM
'Then you would go home.'
(Part of the description of how puberty rites were performed in the past)
(Ablabe_081002_PA)

The perfective has a default deictic temporal reference that depends on the situational aspectual properties of the clause. Actions and punctual events are generally interpreted as occurring prior to the time of speech, as in (48). States are generally interpreted as currently holding at the time of speech. Inchoatives are typically interpreted to have completed in the past and so also yield a present (end) state interpretation, as in (49). These default temporal interpretations can easily be overridden. For instance, actions and punctual events in the perfective can be interpreted as occurring in the speaker's future (51) and states can be interpreted as occurring in the speaker's past, as in (52).

(51) *ɔ̀-lagɔ̀=lɔ* *ɛ̀-dɔ* *gì* *kị-tɔ* *bị-dɔ̀mɛ* *pɔ́*
$C_2$s-evening=DEF $C_2$s.PFV-fall CLM 1p.PFV-cook $C_4$p-thing finish
*te* ***ma-kɔ̀*** *ke-plikpa* *kɔ* ***mà-kpasị̀***
like.this **1s.PFV-take** $C_6$s-book so **1s.PFV-learn**
'In the evening, when we finish cooking, I will take my book and learn.'
(Interview_081015_KA&RE)

(52) *lɛ̌* ***be-zè*** *lɔ̀=ɛ*
and **$C_1$s.PFV-be** there=CM
'And they dwelt there.' (History_081120_WO)

2.7.2.1.2 Progressive

The progressive is used to refer to ongoing actions and situations. The marker is fused with the Set 1 subject agreement prefixes: it consists of the onset of the prefix followed by *ìɪ́-/ị̀ɪ́-* if the vowel of the prefix is i/ị, or *èé-/ɛ̀ɛ́-* otherwise, see (53) and (54).

(53) *o a-sị̀ li-bo = lè* ***lị̀ɪ́-trɛ́*** *l'* *abà* *yà* *na*
oh $C_1$s.PFV-say $C_3$s-matter = DEF **$C_3$s.PROG**-go:LOC $C_3$s on here QM
'Oh, he said, what's going on here?' (Folkstory_110406_QM_00:39)

(54) ***bɛ̀ɛ́-ŋwɛ̀***
**$C_1$p.PROG**-drink
'They were drinking.' (Avatime-history_110905_BB_06:55)

2.7.2.1.3 Habitual

The habitual is marked using the prefix *zɛ̌-/zɛ̌-* and subject agreement markers from Set 3 with an extra-high tone. It is used to refer to actions or situations which are repeated frequently over an extended period of time and are considered usual or predictable, as in examples (55) and (56).

(55) *xé* *bɛ-trɛ* *kù-de = ò* *ɔ-ŋwá* *te*
when $C_1$p.PFV-go $C_5$s-road = DEF INF-weed like.this
***bɪ́-zɛ̌-pɔɪ̀*** *bɪ̀-dɔ̀mɛ* *ŋà*
**$C_1$p-HAB-roast** $C_4$p-thing eat
'When they went to weed the road like that, **they used to roast** food to eat.'
(History_081120_WO)

(56) *kɔ níté* *kʊ́-zɛ̌-wà* *ɪ̀-nyɔ = nɛ̀* *yà*
so how **1p-HAB**-do.work $C_2$p-farm = DEF here
'So this is how we farm here.' (Farming_080909_SO)

2.7.2.1.4 Potential

This mood is used for potential events, these can either be in the future (57) or other events with uncertain epistemic status (58). The marker is fused with the subject agreement prefix. For some noun classes and person/number values it seems to have fused with the Set 1 prefix and for others with Set 3. The fused marker always consists of the onset of the prefix followed by *áà* or *íà/ɪ́à,* see Table 2.10.

(57) ***áà-trɛ*** *ní* *ìgɛ*
**$C_1$s.POT-go** LOC Accra
'He will go to Accra.' (Elic-R_0811191_AB)

(58) ***áà-ze*** *ní* *ìgɛ*
**$C_1$s.POT-be** LOC Accra
'He might be in Accra.' (Elic-R_0811191_AB)

Table 2.10: Form of the fused potential mood marker

| | Set 1 | Set 3 | **Subject Agreement + Potential** |
|---|---|---|---|
| 1s | me-/ma- | mí-/mį́- | **máà-** |
| 1p | ki-/kị- | kí-/kį́ | **kíà-/kį́à-** |
| 2s | wo-/wɔ- | wú-/wṵ́- | **wáà-** |
| 2p | mle-/mlɛ- | mlí-/mlį́- | **mláà-** |
| $C_1$s | e-/a- | i-/ị- | **áà-** |
| $C_1$p | be-/bɛ- | bi-/bị- | **bíà-/bį́à-** |
| $C_2$s | è-/ɛ̀- | ì-/ị̀- | **áà-** |
| $C_2$p | ì-/ị̀- | ì-/ị̀- | **íà-/į́à-** |
| $C_3$s | li-/lị- | li-/lị- | **líà-/lį́à-** |
| $C_3$p | e-/ɛ- | i-/ị- | **áà-** |
| $C_4$s | ki-/kị- | ki-/kị- | **kíà-/kį́à-** |
| $C_4$p | bi-/bị- | bi-/bị- | **bíà-/bį́à-** |
| $C_5$s | ki-/kị- | ki-/kị- | **kíà-/kį́à-** |
| $C_5$p | be-/bɛ- | bi-/bị- | **bíà-/bį́à-** |
| $C_6$s | ke-/kɛ- | ki-/kị- | **kɛ́à-/kɛ́à-** or **kíà-**[5] |
| $C_6$p | ki-/kị- | ki-/kị- | **kíà-/kį́à-** |
| $C_7$ | si-/sị- | si-/sị- | **síà-/sį́à-** |

2.7.2.1.5 Subjunctive

The subjunctive is indicated by the use of subject agreement forms from either Set 2 or 3 depending on the verb. There is no clear pattern as to which verbs take Set 2 and which take Set 3 (Defina 2009).

The subjunctive is used primarily in subordinate clauses; especially for the sentential objects of the verbs *nu* 'be' and *pɛ* 'want' (see (59) and (60)). With *pɛ* 'want', the subject of the subordinate clause must be the same as the subject of the

[5] There is variation between speakers here.

main clause otherwise the perfective will be used instead of the subjunctive, as in (61). It is also used in subordinate clauses which express a reason or motivation for the main clause (62).

(59) *li-nu* *sị̀* ***mí-zè*** *ke-pe = à* *ki-vò*
$C_3$s.PFV-be COMP **1s.SBJV-be** $C_6$s-house = DEF $C_4$s-tomorrow
'I must be home tomorrow.' (Elic-R_081125-3_MM)

(60) *wɔ̀-pɛ* ***wʉ́-trɛ*** *ní* *amekúkúbɔ = ɛ* *mɛ̀*
2s.PFV-want **2s.SBJV**-go LOC cemetery = DEF inside
'You want to go to the cemetery.' (Elic-RS_080922_SO_08:38)

(61) *mà-pɛ* *sị̀* ***wò-se***
1s.PFV-want COMP **2s.PFV-run**
'I want you to run.' (Elic-RS_080905-2_SO)

(62) *lɛ̌* *e-mu* *ku* *ní* *lì-fu = nè* *sì* ***ị́-gʉ́***
then $C_1$s.PFV-climb arrive LOC $C_3$s-house = DEF COMP **$C_1$s.SBJV-pluck**
'Then he climbed the tree in order to pick them' (Pear_081008_HO)

The subjunctive can also be used in main clauses to indicate the speaker thinks the situation should hold, as in (63), or to make a polite request, as in (64).

(63) *kʉ́-trɛ*
1p.SBJV-go
'We should go.' (Wake_080926_WB)

(64) *mí-bu = be* *bàsị́ = yɛ*
1s.SBJV-remove = $C_1$p.OBJ show = $C_1$s.OBJ
'Let me explain it to him.' (Ablabe_081002_PA)

2.7.2.1.6 Imperative

The imperative is signalled by the absence of a subject agreement prefix. It is used to give directives to single addressees, as in (65) and (66). If the speaker wishes to give a directive to more than one addressee she must use the subjunctive as in (67).

(65) *ŋà* *bị-dɛ́yà*
eat $C_4$s-thing:PROX
'Eat this!' (Elic-RS_080903-2_SO)

(66) *zĕ-bí=me* *dzédzé* *là*
REC-ask=1s.OBJ another $C_3$s
'Ask me another one (question).' (Interview_080924_KA-MO)

(67) *mla-ŋà* *bị-dɛ́yà*
2p.SBJV-eat $C_4$s-thing:PROX
'Eat this! (directed to more than one person)' (Elic-RS_080903-2_SO)

#### 2.7.2.2 Non-contrastive aspect and mood markers

##### 2.7.2.2.1 Recurrent

The recurrent is marked by the prefix *zĕ-/zɛ̆-*. It is likely derived from a serial verb construction, probably with the verb *zè* 'be'. It has now grammaticalized to a prefix and participates in vowel harmony with the verb root. There are still some remains of its verbal origins. Avatime speakers above the age of 50 often use *zŏ-/zɔ̆-* in place of *zĕ-/zɛ̆-* in cases where there would have been an *o-/ɔ-* serial verb marker vowel. Speakers under the age of 20 exhibit a different alternation: for them *lĭ-/lị̆-* is interchangeable with *zĕ-/zɛ̆-* in all cases of the recurrent. This suggests a shift to the *lí* 'be.at' verb. Neither of these alternations is possible with the habitual whose marker *zĕ-/zɛ̆-* likely comes from the same origins.

The recurrent is used to refer to situations that are ongoing or repeated over a certain interval of time, as in (68)-(70). Example (70) shows it can occur with different contrastive aspects and moods such as the potential. Its semantic function overlaps with both the progressive and the habitual, see Defina (2009) for a discussion of the differences.

(68) *ɔ́-nɔ̀=ɛ* *gì* ***a-zĕ-da*** *bị-dzyà*
$C_1$s-person=DEF REL **$C_1$s.PFV-REC-sell** $C_4$p-meat
'the person who **sells** the meat' (Dog_081002_PA)

(69) ***ma-zɛ̌-wà*** *à-xwɛ̀ = na* *ní* *cocoa marketing board*
**1s.PFV-REC-do.work** $C_3$p-work = DEF LOC Cocoa Marketing Board
*ní* *koforidua = ɛ*
LOC Koforidua = CM
'I **was working** at the Cocoa Marketing Board in Koforidua.' (Life_081122_AB)

(70) *tɔ* ***kɪ̣à-zɛ̌-ba*** *ke-pe = à*
PURP **1s.POT-REC**-come $C_6$s-work = DEF
'We shall be coming home (often).' (Interview_081015_KA&RE)

2.7.2.2.2 Intentive

The intentive is marked by the prefix *tá-*[6]. Like the potential it is often assumed to have future reference, as in (71). Unlike the potential, its core meaning is one of intention rather than possibility, as can be seen in (72).

(71) *be-do* *sị̀* *yɛ́* ***bɛ-tá-kɔ̀*** *ɛ-wà*
$C_1$s.PFV-say COMP $C_1$s:FOC **$C_1$s.PFV-INT-take** SVM.$C_1$s.PFV-do.work
*kunu = yè*
funeral = DEF
'They said it is him they will use for the funeral.'
(Folkstory_110406_QM_06:24)

(72) ***bɛ-tá-ŋà*** *dɔ̀*
**$C_1$s.PFV-INT-eat** thing
'They were going to eat.' (History_081120_WO)

2.7.2.2.3 Epistemic certainty

Avatime also has a marker of epistemic certainty *nya-*, which is often translated as 'really' or 'certainly' (see (73)). The same marker occurs in Ewe (Ameka 2008) and Tafi (Bobuafor 2013:229). It was likely borrowed from Ewe into both Avatime and Tafi.

[6] *trá-* in the Amedzofe dialect.

(73) *wɔ̀-nya-pɔnị̀=mɛ*
2s.PFV-CERT-help=1s.OBJ
'You really helped me.' (Ablabe_081002_AD_00:43)

#### 2.7.2.3 Periphrastic aspect and modality

Several kinds of periphrastic constructions are used to modify the aspect or modality of Avatime clauses. Like other Kwa languages (Ameka 2006a), but unlike many other serializing languages (Aikhenvald 2006), Avatime does not appear to use serial verb constructions for aspectual or modal functions.

##### 2.7.2.3.1 Phase and ability in non-finite complement constructions

Non-finite complement constructions (Section 2.7.6) are often used to specify the phase or to make a claim about the possibility of a situation. The verb modifying the phase or ability occurs as the main finite verb and takes the rest of the predicate as a non-finite verb phrase complement. If the non-finite verb phrase has an object then the order is reversed, with the object occurring before the verb. An example in shown in (74) the verb *kpese* 'start' used to indicate inception. The verb *hà* 'near, approach' can be used to indicate the prospective, as in (75). The modal verb *tanị̀* 'be able' is used for the ability of an agent to perform an action (76) or a situation's likelihood.

(74) *lɛ̌* *o-bi* *tsyɛ* ***e-kpese*** ***o-se***
then $C_1$s-child also **$C_1$s.PFV-start** **INF-run**
'Then the child too **started running**.' (Expnew06_s2_130805)

(75) *lɛ̌* *ɔ-vlà=ɛ* *a-ga* *xé* ***ị-hà*** ***yɛ***
and $C_1$s-second=DEF $C_1$S.PFV-walk CLM **$C_1$s.SBJV-near** **$C_1$s.obj**
***na=ɛ***
**reach=CM**
'And the second one was walking and **about to reach him**.'
(Contrexp14_s3_120908_13:29)

(76) *xé* ***bị́à-tanì*** ***sɛ*** ***ɔ-ŋwlịmì***
then **$C_1$p.POT-be.able** **$C_7$.OBJ** **INF-write**
'Then **they can write it** (Avatime).' (Language-use_130810)

2.7.2.3.2 Necessity

Necessity is encoded using a fixed phrase *linu sị̀* 'it is that', which takes the main predicate as a clausal complement, as in (77).

(77) *li-nu* *sị̀* *be-hè* *ò-gbe=nò*
$C_3$s.PFV-be COMP $C_1$p.PFV-pull $C_2$s-rope=DEF
'They have to pull the rope.' (Folkstory_110406_QM_06:20)

2.7.2.3.3 Completive particle

There is at least one final particle with an aspectual function: *pɔ́* 'finish', see (78).

(78) *mɛ-hà* *bìtɛ* *pɔ́*
1s.PFV-near do finish
'I've nearly finished.' (Conv-home_100716_AS_07:00)

2.7.3 Negation

Negation is indicated by the use of subject agreement prefixes from Set 2 with an extra-high tone, see (79).

(79) *a-mɔ̀=yɛ* 'He saw her.'
*ɔ́-mɔ̀=yɛ* 'He didn't see her.'

*ki-te* 'We know.'
*kú-te* 'We don't know.'

*bị-pɛ̀* 'It is good.'
*bị́-pɛ̀* 'It is not good.'

The contrastive aspect markers have different forms in the negative. The positive progressive marker *ɛ̌-/ɛ̌-* marker is replaced by *lí-/lị́*, as in (80). This form is likely related to the 'be.at' verb *lị́*.

(80) a. *mèɛ́-ŋà* *blàlị*
1s.PROG-eat plantain
'I am eating plantain.'
b. *mɔ́-lí-ŋà* *blàlị*
1s.NEG-PROG-eat plantain
'I am not eating plantain.' (Elic-RS_080903-2_SO)

The negative habitual is marked by a rising tone on the subject prefix, as in (81).

(81) a. *mí-zɛ̌-ta* *kị-mịmị*
1s-HAB-eat $C_4$s-rice
'I usually eat rice.'
b. *mɔ̌-ta* *kị-mịmị*
1s.NEG.HAB-eat $C_4$s-rice
'I don't usually eat rice.' (Elic-RS_080903-2_SO)

There is no negative form for the potential in the Vane dialect. The negative perfective with the intentive is used instead, even in situations where the intentive would not otherwise be used, such as (82). Funke (1909) and Ford (1971a) report a negation marker *bí-/lí-*, which could be used in conjunction with the potential mood, as in (83). I have found some examples of this marker in the Amedzofe dialect but none in the Vane dialect. It seems this full negation marker is shifting towards a floating tone marker and this process has completed in the Vane dialect. As discussed in Defina (2009), this shift is the likely reason the potential can no longer be negated in the Vane dialect.

(82) *ɔ́-tá-bịtɛ* *pɔ́*
$C_1$s.NEG.PFV-INT-do finish
'He might not have finished it.' (Elic-R_081119-1_AB)

(83) *ɔ̀-bí-â-bịtɛ̀ = bɛ*
$C_1$s-NEG-POT-do = $C_1$p.OBJ
'He will not do it.' (Ford, 1971a: 201)

The subjunctive and imperative can also not be directly negated. Rather the positive subjunctive is used with the prohibitive *kú-/kṵ́-* or *kí-/kṵ́-*, for negations of both subjunctives, as in (84), and imperatives, as in (85). Note that the verb *bìtɛ* 'do' is irregular in the 2nd person singular so that *wo-kú-* reduces to *ú-*, as shown in (86).

(84) *áà-sị = wɔ* *sì* *wụ-kṵ́-kpɛ́* *yɛ* *ò-nugu*
$C_1$s.POT-tell = 2s.OBJ COMP 2s.SBJV-PROH-put:LOC $C_1$s.POS $C_2$s-mouth
'He will tell you that you shouldn't worry him.' (Interview_081120_GE-MM)

(85) *wo-kú-trɛ* *ní* *lị-gba = lɛ̀*
2s.SBJV-PROH-go LOC $C_3$s-building = DEF
'Don't go to the house.' (Lego_081114_AB-WO_3)

(86) *ú-bìtɛ* *bị-dɛ́lɔ̀*
2s.SBJV.PROH-do $C_4$p-thing:DIST
'Don't do that thing.' (Elic-RS_080903-2_SO)

The particle *aní* 'not' can also be used to form negations. It scopes over a particular noun phrase, focused element (87), or complement clause (88).

(87) *pɔ̀* *aní* *kị-mịmị = ɛ́* *kṵ́à-zɔ̌-ta*
but not $C_4$s-rice = DEF:FOC 1p.POT-REC-eat
'But it is not only rice that we shall be eating.' (Greetings_130807_PKD)

(88) *lɛ loso* *xé* *bṵ́à-pɔnɔ̀ = ɛ* *aní* *sị* *ba-zɛ̌-bàsị*
$C_3$s reason if $C_1$p.POT-help = CM not COMP $C_1$p.PFV-REC-show
*suku = ye* *ní* *ba* *ke-pe = à* *kò*
school = DEF LOC $C_1$p.POS $C_6$s-house = DEF just
'So that if they will help us, not that they will go and teach their schools.'
(Language-use_130810)

### 2.7.4 Directionals

There are two directional prefixes which can occur immediately before the verb stem[7]: *ze-/zɛ-* 'itive' and *bá-/bé-* 'ventive', see (89).

(89) a. *a-ze-ku*
$C_1$S.PFV-IT-arrive
'She arrived (there).'

b. *a-bé-ku*
$C_1$S.PFV-VENT-arrive
'She arrived (here).'

c. *èɛ́-bá-ku*
$C_1$S.PROG-VENT-arrive
'She is arriving (there).' (Elic-S_081117-2_MM)

These forms likely came from serial verb constructions with the verbs *za* 'pass' and *ba* 'come'. There are some indications these markers have not completely grammaticalized as prefixes and are still treated as full verbs within a serial verb construction. Firstly, regardless of the ATR value of the directional, the prefixes take –ATR values as would have been triggered by the original verbs. Secondly, both directionals show ATR alternation but only the itive consistently harmonises with the verb root. The alternation for the ventive may be due to the vowel that would have been prefixed to the verb root if the directional and main verb formed a serial verb construction (see Section 3.2.1.2). It seems to have been separated from this though, since younger speakers who tend not to use the serial verb construction vowels also show the alternations in the ventive marker.

### 2.7.5 Comitative

The comitative *-nì/-nị̀* or *-nɔ̀/-nɔ̀* is the only verbal suffix in Avatime. It adds a comitative argument to the predication, as can be seen in (90). The *–nì/-nị̀* form is used before an object, while *–nɔ̀/-nɔ̀* is used at the end of a verb phrase.

---

[7] Note that Ford (1971a) reports these forms occurring before the recurrent marker and sometimes even before the habitual marker. This may be another difference between the Amedzofe and Vane dialects.

(90) a. *a-sɛ̀*
$C_1$s.PFV-leave
'She left.'
b. *bɛ-sɛ̀-nì=yɛ*
$C_1$p.PFV-leave- COM = $C_1$s.OBJ
'They left with her/They took her away.'

It is no longer fully productive. There are several verb + comitative combinations with idiosyncratic meanings, such as *ku* 'arrive' vs. *ku-nì* 'follow'. There are also several cases where only the verb + comitative form remains and the original verb has been lost, such as *panì* 'talk (with/to)' where there is no verb *pa*.

### 2.7.6 Infinitives and nominalization

Verbs in non-finite complements can take one of three forms: they can be unmarked; prefixed with the infinitive prefix *o-/ɔ-*; or prefixed with the noun class 5 singular prefix *ku-/kụ-* resulting in a semi-nominalised form. All three strategies are shown in example (91). There are no clear differences in usage.

(91) *e-kpese* *ò-ni=nò* *tɔ/ɔ-tɔ/kụ-tɔ*
$C_1$s.PFV-start.to $C_2$s-soup = DEF cook/INF-cook/$C_5$s-cook
'She started cooking the soup.'

The semi-nominalised forms cannot function as full nominals. In order to fully nominalise a verb the verb root must be reduplicated. It can then function as the subject of a clause (92), take definite (93) and indefinite (94) articles, be possessed (95), and head relative clauses (96). Nominalised verbs normally occur in noun class 5 but they may occur in other noun classes as in (97).

(92) *anì* ***ku-mu~mu*** *tsyɛ* *ku-lí* *lɛ=mɛ̀*
and **$C_5$s-REDUP~be.tall** also $C_5$s.PFV-be.at $C_3$s = inside
'and there is **drunkenness (lit. tallness)** inside too.'
(Family-problems-task_110316_SO)

(93) *blɔ* *kò* *kɨ́-zɛ̌-pɔnì=yɛ* *ní* *si-deme=se*
1p.SBJ just 1p.HAB-HAB-help = $C_1$s.OBJ LOC $C_7$-avatime = DEF

***ku-kpasị̀~kpasị̀=ɔ*** *mɛ̀*
**$C_5$S-REDUP~learn=DEF** inside
'We help her with **learning** the Avatime language.' (Language-use_130810_BT)

(94) *meɖe 'ku,* *mà-pɛ* *mí-vi* ***ku-vi~vi=itɔ***
please 1s.PFV-want 1s.SBJV-ask **$C_5$S-REDUP~ask=INDEF**
'Please, I would like to ask **a question**.' (Interview-past-CD_081002_WE)

(95) *lɛ̌* *wo* *ègé* *wɔ-tá-bị̀tɛ* *ní* *wo* ***ku-tsì~tsì=o***
and 2s.SBJ what 2s.PFV-INT-do LOC 2s.POS **$C_5$S- REDUP~grow.old=DEF**
*mè*
inside
'And as for you, what will you do when you are grown up?'
(Interview_081001_A)

(96) ***ku-ye~ye*** *kɔ-lɔ̀=ɛ* *gì* *bíà-do* *sị̀*
**$C_5$S-REDUP~kill** $C_5$S-DIST=DEF REL $C_1$P.POT-say COMP
*wáà-pe* *ní* *ku-bi-pɔ* *mɛ̀=ɛ*
2s.POT-suffer LOC $C_5$S-baby-birth inside=CM
That **pain** that they will say you will suffer during childbirth.'
(Adam-and-Eve-story_110409_AB)

(97) *lɛ̌* *ba-tɔ* *be-ple* ***ò-ple~ple=lò***
and $C_1$P.SBJ-INDEF $C_1$P.PFV-descend **$C_2$S-REDUP~descend=DEF**
*ɛ-zɛ̌-dɔ* *ní* *kălètụị̀=à*
SVM.$C_1$P.PFV-IT-move.down LOC lower.side=DEF
'And some descended **downwards** to the lower side.'
(Avatime-history_110905_BB_10:35)

## 2.8 Ideophones and adverbs

### 2.8.1 Ideophones

Following Dingemanse (2011:15), ideophones are defined as "marked words that depict sensory imagery". Structurally, ideophones often contain repetition of syllables or syllable combinations such as *fotsofotso* 'light (weight)', *trátrátrátrá* 'very

neat' and *rìdìdìdìdì* 'continuously'. These syllables or syllable combinations can often be repeated as much as the speaker desires. Another characteristic shared by many ideophones is a long final vowel, as in *hããããã* 'intensely staring' and *blèwùù* 'slowly'. This long vowel can also be extended as long as the speaker wishes. Ideophones can contain syllables ending in a nasal consonant as in *pîm* 'very big', unlike most other words in Avatime. As Dingemanse (2011) observed for Siwu, ideophones are on average longer than other words, which tend to have monosyllabic roots (see Section 2.2.3).

Ideophones can occur in different places in the sentence. They are frequently used adverbially to modify the predicate. For instance the ideophone in (98) provides further specificity to a related verb and the one in (99) modifies the aspect.

(98) *ò-besì=lo* *pɔ̀=ɛ* *ò-nu* *ɔ-ga*
$C_2$S-sheep=DEF CTR=CM $C_2$S.PFV-be $C_1$S-animal
*gì* ***ɛ-hwa*** ***pịtịtịtị***
REL **$C_2$S.PFV-be.white** **ID.white**
'As for the sheep, it is an animal which is very white.'
(Folkstory- chiefsson_110924_PKD)

(99) *bèé-hè* *ò-gbe=nò* ***rìdìdìdìdìdì***
$C_1$P.PROG-pull $C_2$S-rope=DEF **ID.continuously**
*ò-gbe=nò* *e-dzè*
$C_2$S-rope=DEF $C_2$S.PFV-be.long
'They were pulling the rope for a long time, the rope was long.'
(Folkstory_110406_QM)

They can also be used as adjectives (see (100)), or occur independently, outside the structure of the sentence as in the last word of (100) and (101).

(100) *kò* *ɔ-kàtsì* *ɔ́-tɔ* *nì* *ɔ̀-tàmi=nɔ̀* ***hwliyaaa***
just $C_1$S-old.man $C_1$S-INDEF with $C_2$S-beard=DEF **ID.long.curly**
***pịtịtị*** *kò* *a-wɔ́lị* *a-dɔ́* *li-fu=nè*
**ID.white** just $C_1$S.PFV-fall SVM.$C_1$S.PFV-move.from:LOC $C_3$S-sky=DEF
*a-dɔ́* *ke-se=à* ***tîm***
SVM.$C_1$S.PFV-land:LOC $C_6$S-ground=DEF **ID.sound.of.landing**

‘Just then, an old man with an unkempt white beard fell from the sky and landed on the ground *tiḿ*.’ (Folkstory_110406_QM)

(101) *me-dí* *lɛ* *mɛ̀* *haaaaa*
1s.PFV-look:LOC $C_3$s inside **ID.thinking.about.something**
‘I looked into it haaaa.’ (Conv-home_ErA_03:40)

### 2.8.2 Adverbs

Most modifiers of verbs, predicates, and sentences are ideophones, but there is also a small set of non-ideophonic adverbs. Some examples are *tàe* ‘a little’, *kóko* ‘already’, *nyàfɛ* ‘maybe’, and *àbla* ‘now’. Also several temporal expressions, such as the words for today and yesterday, function as adverbials even though they are formally nouns. *Òmonò* ‘today’ has the class 2 singular prefix *ò-* and definite article *=nò*. The concepts of ‘tomorrow’ and ‘yesterday’ are expressed by the same noun, *kivò*, with the class 4 singular prefix *ki-*. When it refers to yesterday, the definite article *=e* is added to form *kivòe* (see (102)).

(102) *wò-dzi* *ì-vɔ́ɪ́* *mɔ* *klɔ̀* ***kivòe***
2s.PFV-buy $C_2$p-eggplant:LOC 1s.POS place **yesterday**
‘Did you buy eggplants from me yesterday? (Conv-street_100720_1)

To talk about a few days ago or a few days from now, the root *de* ‘back’ is added to *kivò*, forming *kivòde*. Similarly, when the day talked about is in the past, the definite article *=e* is added, forming *kivòdeè*.

## 2.9 Compounds

Roots can be compounded to form a single word. Common combinations are noun-noun (103), noun-verb (104), and verb-noun (105). The part of speech is determined by the first element of the compound. In noun-verb compounds, the compound noun can either function as the subject (104a,b) or object (104c) of the verb. In some cases more than two roots are compounded such as in (106), this example is also unusual in that the first noun root occurs with a definite marker.

(103) a. *ɔ-kà-da*
$C_1$s-father-older.sister
'aunt (father's sister)'
b. *ka-kúkɔ-bi*
$C_6$s-chicken-baby
'chick'
c. *ke-se-gù*
$C_6$s-tree-stump
'tree stump'

(104) a. *bè-dè-tsa*
$C_5$p-road-meet
'junction'
b. ɔ-*kà-tsi*
$C_1$s-father-grow.old
'old man'
c. *ɔ-sà-yɔ̀*
$C_1$s-cloth-weave
'weaver'

(105) *fɛ-se* = *a*
lie-ground = DEF
'lie down/sleep'

(106) *kù-ni* = *ò-sí-klɔ̀*
$C_5$s-water = DEF-fetch-place
'water fetching place'

## 2.10 Simple sentences

### 2.10.1 Constituent order

The canonical constituent order of Avatime sentences is shown below in (107).

(107) CLM/LD - foc - sbj - verb - indirect obj - direct obj - oblique - adjunct – FP

The leftmost elements of the sentence are clause linkage markers (CLM) such as *lɛ̌* 'and, then' and left dislocated elements (LD). There can be multiple left dislocated elements and they can either precede or follow the clause linkage marker (for more information see van Putten 2014a; van Putten 2014b). These initial elements are optionally followed by a focus marked element. The lexical subject, if there is one, precedes the verb and the object(s) and/or oblique arguments follow the verb. Any adjuncts follow the arguments and the sentence may end in a final particle (see (108) and (109)).

(108)

| *lɛ̌* | *ba* | *tsyɛ* | *bá-nù=a* | *bɛ-kpasì* | *ba* | *sì-yà=sɛ* |
|---|---|---|---|---|---|---|
| and | $C_1$p.SBJ | ADD | $C_1$p-people=DEF | $C_1$p.PFV-learn | $C_1$p.POS | $C_7$-language=DEF |
| **CLM** | **LD** | | **SBJ** | **VERB** | **OBJ** | |

'Then, they too, the people learnt their language.'

(Avatime-history_110905_BB_03:25)

(109)

| *kɔ* | *ba* | *petee* | *akpɔkpɔ=ɛ* | *kó* | *bɛ̀ɛ́-pɛ* |
|---|---|---|---|---|---|
| so | $C_1$p.SBJ | all | frog=DEF | only:FOC | $C_1$p.PROG-look.for |
| **CLM** | **LD** | | **FOC** | | **VERB** |

| *ní* | *ì-se=le* | *mè* | *te* |
|---|---|---|---|
| LOC | $C_2$p-tree=DEF | inside | like.that |
| **ADJ** | | | **FP** |

'So all of them, were they only looking for the frog in the trees like that?'

(Frog_100719_DQ-PhA)

The only obligatory element in most Avatime sentences is the inflected verb, as in (110). There are however, a few cases where an adjective is used predicatively with no copula verb as in (111), though it is more common to use an explicit copula verb (see Section 2.5.1).

(110) *bɛ̀ɛ́-ŋwɛ̀*

$C_1$p.PROG-drink

'They were drinking.' (Avatime-history_110905_BB_06:55)

(111) *lɛ* *dzyɔ́~dzyɔɔ*
$C_3$S REDUP~tall
'It (the mountain) is very tall.' (Avatime-history_110905_BB_08:26)

If the verb is transitive, the object is usually mentioned explicitly. If the object has recently been mentioned in the preceding discourse, however, it is typically expressed by a pronoun. In some cases, this pronoun can be omitted if the object is recoverable from the context. This is more likely to occur in serial verb constructions. An example can be seen in (112). In the first line, the object 'rice and chicken' is introduced. In the second line it is referred to using a pronoun (*kɛ*). In the third line the food is the direct object of the first verb in the serial verb construction, *kɔ̀* 'take', but no pronoun is used to refer to it.

(112) 1 *bɛ̀ɛ́-tɔ* *kɪ̣-mɪ̣mɪ̣=ɛ̀* *nì* *ɔ̀-kʊ́kɔ=lɔ̀* *kɪ̣-dzyà=ɛ*
$C_1$P.PROG-cook $C_4$S-rice=DEF with $C_2$S-chicken=DEF $C_4$S-meat=DEF
'They are cooking rice with chicken.'
2 *bɛ-gba=kɛ*
$C_1$P.PFV-fry=$C_4$S.OBJ
'They have fried it.'
3 *bɛ-tá-kɔ̀* *bɪ̣tɛ* *jollof*
$C_1$P.PFV-INT-take make jollof
'They will use (it) to make jollof rice.' (Folkstory-1_110409_AB)

### 2.10.2 Transitivity

Avatime has intransitive (113), transitive (114), ditransitive (115), and ambitransitive verbs. Ambitransitive verbs always take the undergoer as their subject when they are used intransitively, as in (116).

(113) *brɛ̀dzyɪ̀mɛ̀=ɛ* *e-tse*
snake(sp)=DEF $C_1$S.PFV-die
'The snake died.' (Avatime-history_110905_BB_09:37)

(114) *be-tá-bu* *sɪ̀-vɪ̀=se* *petee*
$C_1$P.PFV-INT-remove $C_7$-husk=DEF all
'They will remove all the husks.' (Rice-farming_100613_EN_05:51)

(115) *a-zɛ-bàsị = blɔ* *bà-lị = à*
$C_1$s.PFV-IT-show = 1p.OBJ $C_5$p-palm.tree = DEF
'He went to show us the palm trees.' (Conv-ablorme_100715_SO-AS)

(116) a. *a-zɛ-wɔ́lị = yɛ* *kpɛ* *ní* *kù-ni = o* *mè*
$C_1$s.PFV-IT-drop = $C_1$s.OBJ put LOC $C_5$s-water = DEF inside
'He went and dropped him in the water.' (Frog_100719_DQ-PhA)
b. *bá-nɔ̀ = a* *petee* *bɛ-wɔlị̀*
$C_1$p-person = DEF all $C_1$p.PFV-fall
'All the people fell.' (Folkstory_110406_QM_03:16)

Several motion and placement verbs specify an obligatory oblique argument (see also van Putten 2009). For instance, the movement verb *dɔ* 'move from' specifies a source argument (117) and the placement verb *trɔ* 'put on' specifies a goal argument, (118).

(117) *bɛ́-dɔ́* *ní* *ò-dzògbè = lo* *ɔ̀-za = lɔ̀*
$C_1$p.PFV-move.from LOC $C_2$s-desert = DEF $C_2$s-direction = DEF
'They came from the direction of the desert.' (History_081120_WO)

(118) *a-trɔ* *ɔ̀-wlà = lɔ* *ní* *ò-nugu = lò*
$C_1$p.PFV-put.on $C_2$s-hand = DEF LOC $C_2$s-mouth = DEF
'He put his hand on his mouth.' (Famprob_110409_DQ-KX)

The only valency changing morphology in Avatime is the comitative suffix discussed in Section 2.7.5. Avatime has no passive construction. The third person (class 1) plural agreement can be used to refer to subjects which are unknown or which the speaker wishes to background. The object can also be left dislocated to make it more salient. An example can be seen in (119), here the object 'you two' is left dislocated and the summoner, which would be first person plural for the speaker, is backgrounded using the class 1 plural.

(119) (In a traditional tribunal, a number of girls have been summoned for breaking a local law. At some point, two men stand up to speak on behalf of the girls. The chief, who is leading the meeting, asks…)

| *mlɔ* | *tịabà* | *tsyɛ* | *bɛ-kị́ = mlɔ* | *ku-plikpá* | *lɔ̀* |
|---|---|---|---|---|---|
| 2p.OBJ | $C_2$p.two | ADD | $C_1$p.PFV-give = 2p.OBJ | $C_6$p-letter | DIST |

'You two, were you also summoned? (literally: did they (I) give those letters to you two as well?)' (Tribunal_100513-4)

### 2.10.3 Focus

Focus is marked in Avatime using a combination of three marking strategies: 1) the focused element is placed in the clause initial focus position; 2) a floating extra-high tone is attached to the final syllable of the focused element; and 3) the end of the clause is marked with a high boundary tone. For instance, in (120) the first sentence gives the canonical structure with no marked focus, while the second sentence illustrates object focus.

(120) a. *mà-panị̀* *mo-ne-da = a*
1s.PFV-greet 1s.POS:$C_1$s-mother-sister = DEF
'I greeted my aunt.'

b. ***mo-ne-da = á*** *mà-panị̀*
**1s.POS:$C_1$s -mother-sister = DEF:FOC** 1s.PFV-greet
'I greeted [my aunt]$_{FOC}$.' (Elic-foc_100602_SO)

When a verb is focused, a copy of the verb root marked with the prefix *ki-/kị-* occurs in the clause initial focus position while the inflected verb remains in situ. As with other kinds of focus, the clause initial copy of the verb is marked with extra-high tone and there is a clause final high boundary tone, see example (121). In many related languages, focused verbs are nominalized (Ameka 2010). The prefix used with Avatime verbs has the same form as the class 4 singular noun class. The class 5 singular prefix *ku-/kụ-* which is normally used in nominalizations (Section 2.7.6) is not used here. These focused verbs may thus be a different form of nominalisation. For more information on focus and other aspects of information structure in Avatime, see van Putten (2014a).

(121) ***kị-hɔ́*** *bɛ-tá-hɔ = lɔ* *àlɔ́* *bíà-to = lo*
**$C_4$s-grind:FOC** $C_1$p.PFV-INT-grind = $C_2$s.OBJ or $C_1$p.POT-pount = $C_2$s.OBJ
*ní* *kị́-dɛ* *mɛ̀*
LOC $C_4$s-mortar inside
'Do they [grind]$_{FOC}$ it, or do they pound it in a mortar?'
(Illness_100616_SO-DS)

### 2.10.4 Questions

#### 2.10.4.1 Polar questions

Avatime polar questions normally have the same structure as declarative sentences. There does not appear to be a prosodic difference and the ambiguity is typically resolved by context. In order to explicitly mark the sentence as a polar question, a sentence final question marker *na* can be added, as in (122). It is more frequently used in embedded questions than if the question is in the main clause. This marker is often reduced to *a* or even just the high tone.

(122) a. *kofí* *á-yáị* *kè-pli = à*
Kofi $C_1$s.PFV-break $C_6$s-calabash = DEF
'Kofi broke the calabash/Did Kofi break the calabash?'
b. *kofí* *á-yáị* *kè-pli = à* *na*
Kofi $C_1$s.PFV-break $C_6$s-calabash = DEF QM
'Did Kofi break the calabash?' (Elic-qa_100525_SO)

#### 2.10.4.2 Content questions

In content questions, the question word typically occurs sentence initially in the position of focused elements (Section 2.10.1). Like focused elements, question words are followed by a floating extra high tone which attaches to their final syllable (see (123)). Content questions are followed by a low boundary tone, resulting in a falling tone on the last syllable if it isn't low (cf. Ford's (1971a) extra low drop tone).

(123) a. ***ŋyaŋwɛ́*** *á-ta* *á-va = nà*
**who:FOC** $C_1$s.PFV-ate $C_3$p-bean = DEF
'Who ate the beans?'

b. ***egé*** *ɔ́-dzɛ* *a-ŋà*
**what:FOC** $C_1$s-woman $C_1$s.PFV-eat
'What did the woman eat?'

c. ***ège*** ***lɛ*** ***lósó*** *ɔ́-dzɛ* *á-ta* *á-va=nà*
**what** **$C_3$s** **reason:FOC** $C_1$s-woman $C_1$s.PFV-ate $C_3$p-bean=DEF
'Why did the woman eat the beans?' (Elic-QUIS-foc_100714_SO)

The sentence final question marker particle can also be added to content questions for emphasis as in (124).

(124) ***nífɔ́*** *máà-mɔ̀* *ki-bù=ye* ***na***
**where:FOC** 1s.POT-see $C_4$s-honey=DEF **QM**
'Where can I find honey?' (Conv-street_100720-2)

Question words can also occur in situ where they have an 'echo question' interpretation signalling repair. This can be seen in (125) where both speakers B and C respond to speaker A's utterance with an in-situ question word.

(125) A: *mà-ŋwị* *lɔ̀* *kókó* *lo*
1s.PFV-appear there already FP
'I already appeared there.'

B: *wɔ-lí* ***fɔ*** *kóko*
2s.PFV-be.at **where** already
'You were where already?'

C: *wɔ-bìtɛ* ***ège***
2s.PFV-do **what**
'You did what?' (Conv-rice_110411-3-2)

## 2.11 Complex clauses

### 2.11.1 Subordinate clauses

#### 2.11.1.1 Relative clauses

Relative clauses immediately follow their head noun and are introduced by the clause linkage marker *gì*. Subjects, as in (126), objects, as in (127), and adjuncts, as in (128), can all be relativized. Relative clauses are often followed by the clitic *=E* which I call a clause marker (glossed CM). This clitic assimilates in both vowel

height and ATR value to the preceding vowel. It follows several types of subordinate and coordinate clauses and also frequently follows left-dislocated elements. It can be seen in all the relative clause examples (126)-(128).

(126) *ɔ́-nɪ́* *lí-ye* *gì* *a-zɛ̌-gbanɪ̀* *kèdánɔ̀=a=ɛ,*
$C_1$s-person $C_1$s-PROX REL $C_1$s.PFV-REC-lead Avatime.people=DEF.p=CM
*ba* *mɛ́* *e-zè*
$C_1$p.POS inside:FOC $C_1$s.PFV-be
'This person who was leading the Avatime people was from within them.'
(Avatime-history_110905_BB_12:14)

(127) *ò-nipó* *lɔ́-lɔ̀* *gì* *be-tsi=i,*
$C_2$s-river $C_2$s-DIST REL $C_1$p.PFV-block=CM
*kù-ni=o* *kʊ́-lɪ́-kpɛ* *é-ple*
$C_5$s-water=DEF $C_5$s.NEG-PROG-put.in SVM.PROG-descend
'That river which they blocked, the water doesn't flow.'
(Avatime-history_110905_BB_07:18)

(128) *kò* *e-se* *dɔ* *nílɔ* *gì* *e-kpò=e*
then $C_1$s.PFV-run move.from there REL $C_1$s.PFV-hide=CM
'Then he ran out from where he was hiding.' (Folkstory_110406_QM_03:48)

2.11.1.2 Complement clauses

Complement clauses are subordinate clauses that function as arguments of the main clause. In Avatime, complement clauses are typically introduced by the complementizer *sɪ̀*, as in (129). The complement following *sɪ̀* is a full sentence since it is possible for elements to be left dislocated within it (van Putten 2014b).

(129) *be-te* *sɪ̀* *be-nèmi=a* *ba-lɪ́* *níyà*
$C_1$p.PFV-know COMP $C_1$p-sibling=DEF $C_1$p.PFV-be.at here
'They know that their siblings are here.' (Avatime-history_110905_BB_02:51)

As in many related languages (Lord 1993), the Avatime complementizer *sɪ̀* appears to have grammaticalized from the verb *sɪ* 'say/tell'. The verb and

complementizer are frequently used together, as in (130), especially when a recipient argument is expressed. If no recipient is expressed the verb *do* 'say' from Ewe seems to be preferred, as in (131).

(130) *ɔ́-nị̀yɛ* *ɛ̀ɛ́-sị̀=mɛ* *sì* *mí-ze-di=ye*
$C_1$s-someone $C_1$s.PROG-tell=1s.OBJ COMP 1s.SBJV-IT-look.at= $C_1$s. OBJ
'Someone was telling me I should go and see her.' (Conv-funeral_100528-7)

(131) *be-zě-do* *sì* *be-dzì-nì=ye* *kivòde*
$C_1$p.PFV-REC-say COMP $C_1$p.PFV-return-COM= $C_1$s. OBJ day.before.yesterday
'They were saying that they brought her back the day before yesterday.'
(Conv-funeral_100528-7)

Some verbs (*di* 'look at', *kị̀* 'give', *pɛ* 'want', and *sị̀* 'say/tell') can take a complement clause without using the complementizer, see (132)-(135). For these verbs it is always optional to use an explicit complementizer, as in (136). Further research is needed to determine if there is a difference in usage when an explicit complementizer is used.

(132) *a-bé-di* *lì-fìflì=nɛ* *lị̀-hlị̀* *plɛ́*
$C_1$s.PFV-VENT-look.at $C_3$s-k.o.porridge=DEF $C_3$s.PFV-spread put.down:LOC
*ke-se=à*
$C_6$s-ground=DEF
'He noticed that the porridge was scattered on the ground.'
(Folkstory_110406_QM_01:53)

(133) *bɛ-kị̀* *ɔ́-dzɛ* *e-pè* *rìdìdìdìdì*
$C_1$p.PFV-give $C_1$s-woman $C_1$s.PFV-suffer continuously
'They allowed the woman to suffer a long time.' (Midwifery_110901_AB)

(134) *a-pɛ* *yị̀-trɛ* *ní* *níyà* *nì* *níyà* *petee*
$C_1$s.PFV-want $C_1$s.LOG.SBJV-go LOC here and here all
'He wanted to go to both here and here.' (Folkstory_110409_AB)

(135) *a-sị̀* *mị̀-gà*
$C_1$s.PFV-say 1s.SBJV-move
'She said I should come.' (Conv-funeral_100528-7_VB)

(136) *bi-zɛ̌-kị̀* *sì* *ba-dzɔ* *ki-bɔ̀=ɛ*
$C_1$p.HAB-HAB-give COMP $C_1$p.SBJV-donate $C_4$s-money=DEF
'They will make people donate money.' (Funeral_100531_MM-EM-part2)

The above examples show complement clauses are used for expressing jussives (135), causatives (136), cognitive states (129), as well as indirect discourse (131) and perception (132). They are also used for expressing purposives as discussed below in Section 2.11.1.3. Direct discourse is rarely expressed in Avatime but when it is a complement clause is also used. The expression of direct discourse then only differs in the viewpoint on the participants and temporal or spatial location. Direct perception can also be expressed using complement clauses, though the marker *gì* used with relative clauses (Section 2.11.1.1) will sometimes be used in place of *sì*, as in (137). The marker *gì* can only be used with direct, not indirect, perception and seems to indicate a closer connection between the complement and matrix clauses.

(137) *me-nu* *gì* *e-bemì* *tatata* *ní* *lị̀ɔfɔ̀* *tị̀bà*
1s.PFV-hear while $C_1$s.PFV-cry loudly LOC ten $C_2$p.two
'I heard him crying loudly at 12 o'clock.' (Conv-home_100716_ErA)

Complement clauses can also be introduced by the clause linkage marker *xé* 'if/when' when the content of the complement clause is presented as uncertain, as in (138).

(138) *mó-te* *xé* *e-tse*
1s.NEG.PFV-know if $C_1$s.PFV-die
'I don't know if he died.' (Famprob_110401_MeD-BeK-story)

2.11.1.3 Purpose and reason clauses

Purpose clauses can be expressed in one of two ways in Avatime: using the complementizer *sị̀* discussed above (Section 2.11.1.2), as in (139), or using the purposive marker *tɔ*, as in (140).

(139) *lóso* *a-ba* ***sị̀*** *yi-bé-di=blo*
so $C_1$s.PFV-come **COMP** $C_1$s.LOG.SBJV-VEN-look=1p.OBJ
*lị-vlɛ́* *lɛ́-yà* *tete*
$C_3$s-morning $C_3$s-PROX like.that
'So she has come to see us this morning.' (Avopa-meeting_100512-1_QM)

(140) *tsyiami* *wáà-nu* ***tɔ*** *blɔ* *petee* *kíà-nu*
spokesman 2s.POT-hear **PURP** 1p.SBJ all 1p.POT-hear
'Spokesman hear, so that all of us will hear.' (Language-use_130810_MgA)

Reason clauses begin with the phrase *lese linu sị̀* or in short *lese sị̀,* as in (141). The word *lese* is likely a contraction of the class 3 singular pronoun *lɛ* and *ese* 'under'. As *nu* is the identificational copula, the long phrase can be translated literally as 'under it is that ...'.

(141) *be-se* *trɛ ní* *ke-pe=a* *mɛ̀* ***lese sị̀*** *ó-nyimemi=yè*
$C_1$p.PFV-run go LOC $C_6$s-house=DEF inside **because** $C_1$s-young.man=DEF
*tole* *a-xwa=ba*
$C_1$s.one $C_1$s.PFV-call= $C_1$p.OBJ
'They ran back to the house because one young man called them'
(FinSto_100614_WE)

2.11.1.4 Temporal and conditional clausal adjuncts

The clause linkage markers *gì* and *xé* are used to introduce clausal adjuncts such as temporal and conditional clauses.

Temporal clauses marked with *gì* are either simultaneous with or prior to the event in the main clause. They typically occur before the main clause but can also follow it. Temporal clauses introduced with *gì* often end with the clause marker introduced in Section 2.11.1.1.

(142) ***gì*** *ó-dí* *dzɛ̀=**ɛ*** *ɛ̀ɛ́-sa* *a-kpɛ=la*
**CLM** $C_1$S.PFV-sit again=**CM** $C_1$S.PROG-hit $C_3$p-hand=DEF
'When he sat down again, he was clapping his hands.'
(Maus-drum_100709_Mia-DQ)

Temporal clauses introduced with *xé* can either be interpreted like those introduced with *gì*, as in (143), or have a more specific 'before' interpretation, as in (144).

(143) *xé* *kʊ̀ʊ́-sɔ* *pɔ́=ɛ* *kɪ-tá-halɪ̀* *sɪ̀-wa=sɛ̀*
CLM 1p.PROG-sow finish=CM 1p.PFV-INT-gather $C_7$-weed=DEF
*petee* *ní* *ɔ̀-nyɔ=nɔ̀* *mɛ̀*
all LOC $C_2$S-farm=DEF inside
'When we finish sowing, we gather all the weeds from the farm.'
(Rice-farming_100613_EN_040)

(144) *blɔ* *kèdánà* *kʊ́-tá-tanɪ̀* *kunu=yè* *ɔ-wa*
1p.SBJ Avatime.people 1p.NEG.PFV-INT-be.able funeral=DEF INF-do
*xé* *kʊ́à-ŋà* *à-mʊ=nà*
CLM 1p.POT-eat $C_3$p-rice=DEF
'We Avatime people cannot perform the funeral rites before we celebrate the rice festival.' (Chiefs-meeting_100610-03)

This 'before' meaning can be made more explicit using the phrase *xé able ke* literally 'before now the same', as in (145). When a temporal clause introduced with *xé* is used with a 'before' meaning, it tends to occur after the main clause. When it is used with a more general simultaneous or possibly preceding interpretation, it tends to occur before the main clause, as in (143).

(145) *mɛ* *mi-vi* *li-boétɔ* *xé* *ablé* *ke*
1s.SBJ 1s.SUBJ-ask $C_3$S-matter:INDEF CLM now same
*wáà-bɪ̀tɛ* *bɪ-dɛ́yà*
2s.POT-do $C_4$p-thing:PROX
'I would like to ask something before you do this thing.'
(Chiefs-meeting_100610-03)

Conditional clauses are introduced either with *xé*, as in (146), or with *xé gì*, as in (147). They typically occur before the main clause, which may begin with *kɔ* 'then'. Conditional clauses typically end in the clause marker, as does the main clause if it starts with *kɔ* 'then'. Note that it can be difficult to draw a line between a temporal and a conditional interpretation as there is a large amount of overlap in the idea 'when x then y' and 'if x then y'. Often it is not, in fact, necessary to draw such a distinction.

(146) *xé* *be-tsyí* *sụ* *te*
CLM $C_1$p.PFV-turn hang like.that
*kɔ* *bɛ̀ɛ́-ŋwɛ̀* *kù-ni=o=è*
then $C_1$p.PROG-drink $C_5$s-water=DEF=CM
'If they hang their heads like that, then they are drinking water.'
(Rice-farming_100613_EN_04:13)

(147) *xé* *gì* *a-zɛ-bàsị=blɔ* *bà-lị=à=ɛ*
CLM CLM $C_1$s.PFV-IT-show=1p.OBJ $C_5$p-palm.tree=DEF=CM
*kɔ* *ki-bu* *wa* *sụ=ị*
then 1p.PFV-remove $C_5$p side=CM
'If he shows us the palm trees, then we'll clear (the bush) around them.'
(Conv-ablorme_100715_SO-AS)

### 2.11.2 Coordination

Avatime clauses can be conjoined using a number of different conjunctions. Some of these, like *lɛ̌, kɔ, pɔ̀*, and *àló,* are used primarily for coordinating clauses, while others, such as *xé* and *gì* are also used with subordinate clauses. Below I list the attested situations for each conjunction. Further research would be needed to determine the ranges of syntactic structures compatible with each form and the relationships between these syntactic structures, the forms, and their functions.

#### 2.11.2.1 lɛ̌

The coordinator *lɛ̌* is used to connect clauses describing events which have already occurred or are currently occurring. The two events may follow each other, as in (148), or occur simultaneously, as in (149). This construction is used in one of the experiment conditions in Chapter 5.

(148) ***lɛ̌*** *ɔ-nị̀vɔ=ɛ* *a-trɛ* ***lɛ̌*** *bɛ-vɔ* *lì-fiflì=nɛ*
**and** $C_1$s-child=DEF $C_1$s.PFV-go **and** $C_1$p.PFV-mould $C_3$s-t.o.porride=DEF
***lɛ̌*** *bɛ-kɔ* *ɛ-kí=yɛ*
**and** $C_1$p.PFV-take SVM.$C_1$p.PFV-give= $C_1$s.OBJ
'And when the child went, they moulded the porridge and gave (it) to him.'
(Folkstory_110406_QM_02:16)

(149) (Description of a video in which two events happen simultaneously)
*ɔ-nị̀vɔ=ɛ* *èé-se*
$C_1$s-child=DEF $C_1$s.PROG-run
***lɛ̌*** *ɔ-kàtsi=e* *èɛ́-gà=**ɛ***
**and** $C_1$s-old.man=DEF $C_1$s.PROG-walk=**CM**
'The child is running and the old man is walking.' (Expsg06_05runb)

Clauses that are coordinated with *lɛ̌* often end in the clause marker (149), which is also used with relative clauses and clausal adjuncts preceding the main clause (see Section 2.11.1).

*Lɛ̌* does not always conjoin two clauses; it is also frequently used sentence initially to indicate the continuation of a story, as in (148).

#### 2.11.2.2 kɔ

The coordinator *kɔ* is used to connect clauses describing events not known to have happened yet (see (150)) or generic events. As with *lɛ̌*, the temporal relation between the two clauses is unspecified. Clauses that start with *kɔ* also often end in the clause marker.

(150) (The speaker is discussing plans for an event to be held later in the year)
*kui-tè* *sị* *bíà-kpese* *dɔ̀mɛ* *ní* *gbàdzɛmɛ̀*
1p.PFV-know COMP $C_1$p.POT-start thing LOC Gbadzeme
***kɔ*** *bɛ-bá* *babiakpa=**ɛ***
**and** $C_1$s.pfv-come:LOC Biakpa=CM
'We know they will start the thing in Gbadzeme and come to Biakpa.'
(Chiefs-meeting_100619-03)

Like *lɛ̌*, *kɔ* can also be used to indicate continuation rather than coordinating two clauses. This typically occurs in descriptions of planned events, instructions and procedural descriptions, for instance, in line 5 of (151), where the speaker returns to describing how to perform the ritual after a small aside.

(151) (A woman describes a ritual that used to be performed for weddings)

1 ***kɔ*** *ɔ̀-klị̀pɔ̀ = lo* *kɔ,* *bɛ-tá-vu* *wlo-nì = wó*
**and** $C_2$s-witness = DEF CTR, $C_1$p.PFV-INT-hold bathe-COM = 2s.OBJ:LOC
se mɔ̀mɔ̀mɔ̀mɔ̀
$C_7$ ID.very.well
'As for the witness, they would hold you and bathe in it (the mix of clay and water) very well.'

2 *ŋwasị̀* *sì* *wu-bemì* *tsyɛ* *wá-mɔ̀* *e-bemì*
be.like COMP 2s.SBJV-cry ADD 2s.NEG.PFV-see SVM.2s.PFV-cry
'You feel like crying but you cannot cry.'

3 (laughs)

4 *ì-klị̀pɔ̀* *ɛ̀-tɔ* *kɔ* *ì-kume = me*
$C_2$p-witness $C_2$p-INDEF CTR $C_2$p.PFV-hurt = 1s.OBJ
'Some roles I played as a witness were painful (referring to the bathing with mud)'

5 ***kɔ*** *mlɛ́-sɛ́ = ɛ* *kɔ* *mlɛ-trɛ́* *àmèdzòfɛ* *mɛ́*
**and** 2p.PFV-leave = CM CTR 2p.PFV-go:LOC Amedzofe inside
*kálà = ɛ*
downstream = CM
'Then you would leave and you would go to the downstream side of Amedzofe.' (Ablabe_081002_AD-YD)

When *kɔ* is used in a non-conjoining way with events that are known to have happened, it signals the start of a new episode or topic, as in line 2 of (152).

(152) 1 *lɛ̌* *kà-tụ̀kpa = a* *a-wɔ̀* *lị̀-ŋwàfụ̀ = nɛ* *mɛ̀*
and $C_6$s-male.goat = DEF $C_1$s.PFV-remain[8] $C_3$s-forest = DEF inside

[8]The goat is personified in this story and so takes class 1 agreement.

*xé* *èɛ́-sɔ́* *yɛ* *ɔ̀-nyɔ = nɔ* *mɛ̀*
and $C_1$s.PROG-hoe $C_1$s.POS $C_2$s-farm = DEF inside
'And the goat remained in the forest and he was hoeing his farm.'

2 ***kɔ*** *e-wè = la* *gì* *bɛ-trɔ* *kɪ́* *ɔ-kàtsi = e*
**and** $C_3$p-day = DEF REL $C_1$p.PFV-put.on give $C_1$s-old.man = DEF
*kunu = yè* *e-wè = la* *ɛ-na = ɛ*
funeral = DEF $C_3$p-day = DEF $C_3$p.pfv-reach = CM
'So the day they set for the old man's funeral, the day has arrived.'
(Folkstory_110406_QM_06:12)

#### 2.11.2.3 xé

The connector *xé* is not only used to introduce clausal adjuncts (Section 2.11.1.4), it can also be used to coordinate clauses. It is often used to coordinate clauses within larger subordinate clauses, as in (153), where the larger subordinate clause is indicated with brackets.

(153) *le* *de* *[xé* *gì* *a-mụ = nà* *èɛ́-pɔ̀*
$C_3$s back [CLM CLM $C_3$p-rice = DEF $C_3$p.PROG-give.birth
***xé*** *e-kpese* *kù-ni = o* *ŋwɛ̀]*
**CLM** $C_3$p.PFV-remain $C_5$s-water = DEF drink]
*kɔ* *séfofo = a* *beé-tsyí* *sụ = ị*
and flower = DEF.p $C_1$p.PROG-turn hang = CM
'After that, if the rice is germinating **and** is starting to drink water, then the flowers will be hanging their heads.' (Rice-farming_100613_EN_04:05)

It can also be used much like *lɛ̌*. There seems to be a tendency to use *xé* to connect more closely related events, as in (154). There is, however, no clear line between contexts where *lɛ̌* is used and contexts where *xé* is used, as shown in (155) and (156) where the examples in (a) show two clauses joined by *xé* and those in (b) show similar clauses joined by *lɛ̌*.

(154) *rrr* *si-se = se* *mè* *kéé* *kù-ni = ò* *kụ̀ɪ́-gà*
continuously $C_7$-mud = DEF inside really:FOC $C_5$s-water = DEF $C_5$s.PROG.move

***xé*** *kui-zě-dò* *ní* *kălètụịa* *ní* *ɔ̀-kụ = tɔ*
**CLM** $C_5$s.HAB-HAB-exit LOC downside LOC $C_2$s-place = INDEF
'The water is running deep in the ground **and** coming out at some lower place.' (Avatime-history_110905_BB _07:47)

(155) a. *ńte* *mè* *sị̀* *ba* *bɛ-bla* *ba* *dɔ̀mɛ* *petee*
LOC.like.that inside COMP $C_1$p.SBJ $C_1$p.PFV-pack $C_1$p.POS thing all
***xé*** *bɛ-sɛ̀*
**CLM** $C_1$p.PFV-leave
'So, they packed all their things **and** left.' (Folkstory_110406_QM_04:41)

b. *gì* *a-halị̀* *lị̀-fị̀flị̀ = nɛ,*
CLM $C_1$s.PFV-gather $C_3$s-t.o.porridge = DEF
*a-kɔ̀* *a-kpɛ́* *ò-gudo = lo* *mè* *kò*
$C_1$s.PFV-take SVM.$C_1$s.PFV-put.in:LOC $C_2$s-bag = DEF inside just
***lɛ̌*** *a-kɔ̀* *a-tsịnị́* *ke-pe = à = ɛ*
**and** $C_1$s.PFV-take SVM.$C_1$s.PFV-send:LOC $C_6$s-house = DEF = CM
'When he collected the porridge, he put it in the bag and took it home.' (Folkstory_110406_QM_01:57)

(156) a. *bɛ-kɔ̀* *kà-tụ̀kpa = a* ***xé*** *bɛ-kɔ̀* *e-ye*
$C_1$p.PFV-take $C_6$s-male.goat = DEF **CLM** $C_1$p.PFV-take SVM.$C_1$p.PFV-kill
'They took the goat and killed it.' (Folkstory_110406_QM_07:22)

b. *lɛ̌* *be-vù* *ɔ́-dzɛ = ɛ* ***lɛ̌*** *be-ye = e*
and $C_1$p.PFV-hold $C_1$s-women = DEF **and** $C_1$p.PFV-kill = $C_1$p.OBJ
'Then they caught the woman and killed her.' (Avatime-history_110905_BB _15:14)

Like *lɛ̌*, *xé* can be used at the beginning of a sentence to signal the continuation of a story, and clauses beginning with *xé* may end in the clause marker.

#### 2.11.2.1 gì

The marker *gì* is most commonly used to introduce relative clauses, complement clauses, or clausal adjuncts as discussed in Sections 2.11.1.1, 2.11.1.2 and 2.11.1.4. Similar to *xé*, this marker can also be used to coordinate two clauses, as in (157). This coordinating use of *gì* is quite rare though and further research is required to

determine whether they are in fact instances of coordination and if so how they differ from coordinations marked with other connectors.

(157) *e-ble=be* *ní* *lị-klʊ̀ụ̀=lɛ* *mɛ̀*
$C_1$S.PFV-unwrap=$C_4$P.OBJ LOC $C_3$S-package=DEF inside
***gì*** *á-kɔ* *a-kpɛ́* *ò-nugu=lo* *mè*
**CLM** $C_1$S.PFV-take SVM.$C_1$S.PFV-put.in:LOC $C_2$S-mouth=DEF inside
'He unwrapped it and put it in his mouth.' (Contrexp09_s2_120906)

2.11.2.2 pɔ̀ 'but'

The coordinator *pɔ̀* 'but' is used to indicate an adversative relation between two clauses, as in (158).

(158) (Description of a person with a lazy eye believed to be caused by fluid entering the eye at birth)
*èé-di=wɔ* *pɔ̀* *èé-di* *ɔ̀-kɔ* *dzedze*
$C_1$S.PROG-look.at=2S.OBJ but $C_1$S.PROG-look.at $C_6$S-place other
'He is looking at you, but he is looking at another place.'
(Midwifery_110901_AB)

2.11.2.3 àló and pʊtɔ̀ 'or'

Disjunction can be marked using *àló*, or less commonly *putɔ̀*. These markers can be used with two clauses, as in (159), or with two nominals, as in (160).

(159) *kɔ* *bɛ-kị=wɔ* *wɔ-fɛ-sa* *àló* *wo-dí*
so $C_1$P.PFV-give=2S.OBJ 2S.PFV-lie-ground or 2S.PFV-sit
'They'll make you lie down or sit up.' (Midwifery_110901_AB)

(160) (One speaker mentions that the leaves from a certain plant can be used to cure a disease. The other speaker interrupts and asks…)
*à-wɔ̀wɔ̀=la* *pʊtɔ̀* *à-kpakpa=là*
$C_3$P-fresh=DEF or $C_3$P-dry=DEF
'Fresh ones or dry ones?' (Illness_100616_SO-DS)

## 2.12 Conclusion

This chapter has provided an overview of the grammar of Avatime, with a clear exception: serial verb constructions. These constructions which are neither coordinate nor subordinate, and which form the major focus of the present thesis, will be dealt with in detail in the following chapter.

# 3 Serial verb constructions and their subtypes in Avatime

A version of this chapter appears as:
Defina, Rebecca. 2016. Serial verb constructions and their subtypes in Avatime. *Studies in Language* 40(3). 648-680.

## 3.1 Introduction

This chapter presents a description of Avatime serial verb constructions (SVCs), their properties, functions, and subtypes. Such a description is not only necessary background for the present thesis, but also of wider typological interest due to Avatime's status as an agglutinating member of the Ghana-Togo Mountain branch of the otherwise typically isolating Kwa language family. It is also notable for the unusual system of truncated agreement markers and the distinction between nuclear and core SVC subtypes not typically reported for West African languages.

SVCs were first described among the Kwa languages Akan (Christaller 1875) and Ewe (Westermann 1907). Kwa languages have continued to feature prominently in SVC research since (e.g. Aboh 2009; Baker 1989; Bamgbose 1974; Collins 1997; Déchaine 1993; Lord 1993), particularly in the characterization of the West African serializing language type (Aikhenvald 2006; Ameka 2003; Dimmendaal 2001).

The Ghana-Togo Mountain languages are a subgroup of Kwa noted for their typological differences (Heine 1968). These differences include a greater use of verbal morphology and, according to Dimmendaal (2001), a corresponding paucity of SVCs. However, more recent descriptions have reported frequently occurring SVCs (e.g. Ameka 2003; Ameka 2009; Bobuafor 2013; Dorvlo 2008; van Putten 2009) suggesting the reported lack of SVCs was due to the previously limited state of description rather than their use of verbal morphology. Nevertheless, Ghana-Togo Mountain languages are still mentioned relatively scarcely in discussions of SVCs and it is an open question how well they fit the West African serializing prototype which has been primarily based on their isolating Kwa relatives. The present chapter compares Avatime SVCs with those found in other, more prototypical, West African languages. In so doing, it finds many similarities but also many differences. Similarities are also noted between Avatime SVCs and those in other less typical West African serializing languages, such as Isu (Kießling 2011), as well as languages from further afield such as Oceania and East Asia.

### 3.1.1 Defining serial verb constructions

Before beginning the description of SVCs in Avatime, it is necessary to define what I mean by an SVC. As mentioned in Section 1.4, many definitions for SVCs have been proposed over the last decades (e.g. Aikhenvald 2006; Baker 1989; Collins 1997; Comrie 1995; Déchaine 1993; Durie 1997; Foley & Olson 1985; Foley 1997; Haspelmath 2016; Lord 1993; Sebba 1987; Seuren 1991). Due to increasing evidence for variation among SVCs, there has been a shift towards identifying SVCs by resemblance to a prototype rather than by a list of strict necessary and sufficient conditions (Aikhenvald 2006; Durie 1997; Foley 2010a; van Staden & Reesink 2008; Senft 2008). This fits with a growing shift towards the use of many variables to identify prototypical or canonical constructions as in canonical typology (e.g. Brown, Chumakina & Corbett 2012). A good list of the properties believed to be prototypical of SVCs is provided by Durie (1997) and Aikhenvald (2006):

i) They consist of a sequence of two or more verbs which function independently as verbs in monoverbal clauses.
ii) They are monoclausal, with all the intonational properties of a monoverbal clause.
iii) There is one tense, aspect, mood and polarity value that is shared by all verbs. This is normally marked on one verb but is sometimes marked on all.
iv) There are no markers of subordination or coordination.
v) The verbs share at least one core argument.
vi) There is only one grammatical subject.
vii) The construction refers to a single event.

Two of these properties – intonation and referring to a single event – have been the subject of much theoretical concern (e.g. Bisang 2009; Crowley 2002; Foley 2010a; Himmelmann 2013; Senft 2008). The issues centre on how to determine whether or not a construction meets the criterion. However, I believe the issue is more basic and lies in the fact that neither intonation nor event structure are morpho-syntactic properties. Since SVCs are morpho-syntactic constructions, they should be identified solely by their morpho-syntactic properties. It is interesting to consider how intonation and event structures relate to morpho-syntactic construction types such as SVCs, but for this to be done non-circularly they cannot

be involved in the identification of these constructions. Constructions must first be defined for individual languages based on specific language internal morpho-syntactic criteria. Once there is a clearly defined construction type within a language, its properties can be compared with the typical cross-linguistic properties of various construction types and if there is a sufficiently good match it can be identified as an instance of that construction type. Only then may questions regarding how the construction relates to intonation or event structure be evaluated.

## 3.2 Serial verb constructions in Avatime

Avatime makes frequent use of a construction closely matching the crosslinguistic prototype for SVCs. The construction combines two or more verbs within a single clause, with no markers of subordination or coordination. The construction consists of a single clause since illocutionary force and negation necessarily scope over all verbs. The subject must be a shared argument of all verbs, and other arguments may optionally be shared. Two verbs – as in examples (1) and (2) – is the most common, and four verbs – as in example (3) – is the longest attested sequence in my corpus.

(1) *bị-lị̀la* *kú* *li-fu=nè*
$C_4$p.PFV-vanish enter:LOC $C_3$s-sky=DEF
'They vanished into the sky.' (Folkstory_110406_QM_01:48)

(2) *kɔ* *be-bu=ye* *plɛ* *ke-sà*
so $C_1$p.PFV-remove=$C_1$s.OBJ put.down $C_6$s-ground
'So they put him (the baby) down on the ground.'
(Midwifery_110901_AB_04:35)

(3) *ma-kɔ̀=bɛ* *nywa* *kpɛ* *kị́* *ɔ-ka-tsi=e*
$C_1$s.PFV-take=$C_4$p.OBJ throw put give $C_1$s-father-old=DEF
'I threw it to the old man.' (Elicitation_110822_SO)

The fact that these constructions consist of a single clause is sufficient to distinguish them from most other multiverbal constructions in Avatime. These tend to consist of multiple clauses where each verb is fully and independently inflected for subject, aspect, and mood and may also be independently marked for negation

and illocutionary force, for example the causative in (4), see Section 2.11 for more information.

(4) *kɔ bɛ-kị = wɔ wɔ-fɛ = se = a àló wo-dí*
so $C_1$p.PFV-give = 2s.OBJ 2s.PFV-lie = ground = DEF or 2s.PFV-sit
'So they make you lie down or sit.' (Midwifery_110901_AB_025)

The only other constructions with multiple verbs in a single clause are the non-finite subordinate constructions (Section 2.7.2.3.1). In these constructions, a verb – typically a phasal or modal verb such as kpese 'start' or tanị̀ 'be able' – takes a non-finite verb phrase as a complement. The second verb can take one of three forms: bare verb stem (5a), prefixed with the non-finite marker o-/ɔ- (5b), or prefixed with the Class 5 singular noun class prefix ku-/kụ- commonly used with deverbal nouns (5c). If the second verb takes an object it is fronted and occurs before the second verb. There are no clear differences in usage between the three forms and even though the deverbal noun class is used in some cases, the verbs are not fully nominalized. In order to fully nominalise a verb the verb root must be reduplicated.

(5) a. *e-kpese ò-ni = nò tɔ*
$C_1$s.SBJ.PFV-start $C_2$s-soup = DEF cook
'She started to cook the soup.'
b. *e-kpese ò-ni = nò ɔ-tɔ*
$C_1$s.SBJ.PFV-start $C_2$s-soup = DEF INF-cook
'She started to cook the soup.'
c. *e-kpese ò-ni = nò kụ-tɔ*
$C_1$s.SBJ.PFV-start $C_2$s-soup = DEF $C_5$s-cook
'She started to cook the soup.'

While the fully inflected first verb in a non-finite complement construction must be a single simple verb, the second position may be filled by a complex predicate such as another non-finite subordinate construction, as in (6), or an SVC, as in (7).

(6) *èé-kpese* *tanị̀* *tì*
$C_1$s.PROG-start be.able crawl
'He is starting to be able to crawl.'

(7) *be-kpese* *ba-wa=tɔ* *tsrɛ* *kʊ́=wɛ*
$C_1$p.PFV-start $C_5$p-medicine=INDEF change give=2s.OBJ
'They start to change medicines for you.' (Midwifery_110901_AB_08:47)

There are formal and functional similarities between SVCs and non-finite subordinate constructions in Avatime: when the verb is not marked and there is no overt object, they look exactly like an SVC; and they are used for functions expressed using SVCs cross-linguistically (Aikhenvald 2006). Non-finite subordinate constructions and SVCs are, however, clearly distinct construction types in Avatime with different properties. The two verbs in non-finite subordinate constructions are always in a predicate argument relation and the first verb typically modifies the aspect or mood of the situation described by the second verb. In contrast, the verbs in Avatime SVCs are never in a predicate argument relation. They are also distinguished by the different marking possibilities and position of the object for the subsequent verb. In the rest of this section, I describe the morpho-syntactic properties of Avatime SVCs in more detail with particular focus on the marking possibilities of subsequent verbs and argument sharing.

### 3.2.1 Inflection within SVCs

The first verb in an Avatime SVC is fully inflected for subject agreement, negation, aspect, and mood. Subsequent verbs are often not marked as in (1)-(3). They can also optionally be marked by a special reduced agreement prefix consisting of a single vowel and tone. These reduced agreement markers are only used with subsequent verbs in SVCs and are one of the features which distinguish SVCs from other multiverbal constructions in Avatime. I refer to them as serial verb markers (glossed as SVM).

If the SVC is in the perfective, then the serial verb markers are reduced forms of the normal subject agreement prefixes: the initial consonant, if any, is elided leaving the vowel and its associated tone, see examples (8)-(11) and Table 1. The use of agreement prefixes on subsequent verbs in SVCs is quite common among Ghana-Togo Mountain languages such as Likpe (Ameka 2003). The elision of the initial

consonant of subject agreement prefixes in certain syntactic environments, often including SVCs, is also quite common among Ghana-Togo Mountain languages, e.g. Siwu (Kropp Dakubu & Ford 1988) and Tafi (Bobuafor 2013). However, Avatime is the only one so far reported where these truncated subject agreement markers are restricted to subsequent verbs in SVCs. Outside of the Ghana-Togo Mountain languages, truncated subject agreement markers in SVCs have been reported in a few languages, such as Koṇḍa (Trans-New Guinea) and Bislama (Creole) (Aikhenvald 2006:41).

(8) a. ***mè****-se* ***è****-mu* *li-to = lè*
**1s.PFV**-run **SVM.1s.PFV**-climb $C_3$s-mountain = DEF
'I ran up the mountain.'

b. ***wò****-se* ***ò****-mu* *li-to = lè*
**2s.PFV**-run **SVM.2s.PFV**-climb $C_3$s-mountain = DEF
'You ran up the mountain.'

c. ***ki****-se* ***i****-mu* *li-to = lè*
**1p.PFV**-run **SVM.1p.PFV**-climb $C_3$s-mountain = DEF
'We ran up the mountain.'

(9) *rrrrrrr* ***bɛ****-trɛ* ***e****-mu*
IDEO:continuously **$C_1$p.PFV**-go **SVM.$C_1$p.PFV**-climb
'They were going up for a long time.' (Avatime-history_110905_BB_01:57)

(10) *ńte* *mè* *sị̀* *ka-pà = tɔ*
LOC:like.that inside COMP $C_6$s-part = INDEF
***ke****-mu* ***e****-ku* *li-to = lè* *abà*
**$C_6$s.PFV**-climb **SVM.$C_6$s.PFV**-enter $C_3$s-mountain = DEF top
'In that way one group climbed onto the mountain.'
(Avatime-history_110905_BB_10:28)

(11) *ba* *li-bo = le* ***lị́****-dra* ***ɟ́****-kɟ́ = ba*
$C_1$p.SBJ $C_3$s-matter = DEF **$C_3$s.NEG.PFV**-be.clear **SVM.$C_3$s.NEG.PFV**-give = $C_1$p.OBJ
'They say they are not clear with the matter.' (Folkstory_110406_QM_04:12)

If the SVC is not in the perfective, the serial verb marker has a fixed form and does not agree with the subject. If the SVC is in the potential mood, the serial verb marker is either o-/ɔ- (example (12)) or e-/ɛ- (example (13)). The choice of form varies between speakers. Individuals are very consistent in which form they use and the variation does not seem important to Avatime speakers, nor does it correlate with any obvious factors such as dialect, gender, or age.

Table 3.1: Subject agreement prefixes in SVCs in the perfective. Pairs refer to vowel harmony pairs. Prefixes with consistently low and extra-high tones are marked. The tone of other prefixes varies with the tone of the verb root as well as the tone of the antecedent's noun class prefix.

| | Positive | | Negative | |
|---|---|---|---|---|
| | Full | SVM | Full | SVM |
| 1s | me-/ma- | e-/a- | mó-/mɔ́- | ó-/ɔ́- |
| 1p | ki-/kị- | i-/ị- | kú-/kṵ- | ú-/ṵ- |
| 2s | wo-/wɔ- | o-/ɔ- | wó-/wɔ́- | ó-/ɔ́- |
| 2p | mle-/mlɛ- | e-/ɛ- | mlá- | á- |
| $C_1$s | e-/a- | e-/a- | ó-/ɔ́- | ó-/ɔ́- |
| $C_1$p | be-/bɛ- | e-/ɛ- | á- | á- |
| $C_2$s | è-/ɛ̀- | è-/ɛ̀- | ó-/ɔ́- | ó-/ɔ́- |
| $C_2$p | ì-/ị̀- | ì-/ị̀- | í-/ị́- | í-/ị́- |
| $C_3$s | li-/lị- | i-/ị- | lí-/lị́- | í-/ị́- |
| $C_3$p | e-/ɛ- | e-/ɛ- | á- | á- |
| $C_4$s | ki-/kị- | i-/ị- | kí-/kị́- | í-/ị́- |
| $C_4$p | bi-/bị- | i-/ị- | bí-/bị́- | í-/ị́- |
| $C_5$s | ki-/kị- | i-/ị- | kú-/kṵ- | ú-/ṵ- |
| $C_5$p | be-/bɛ- | e-/ɛ- | bá- | á- |
| $C_6$s | ke-/kɛ- | e-/ɛ- | ká- | á- |
| $C_6$p | ki-/kị- | i-/ị- | kú-/kṵ- | ú-/ṵ- |
| $C_7$ | si-/sị- | i-/ị- | sí-/sị́- | í-/ị́- |

(12) a. ***máà****-se* ***ɔ****-sɛ̀*
**1s.POT**-run **SVM.POT**-leave
'I will run away.'

b. ***kíà**-se* ***ɔ**-sɛ̀*
**1p.POT**-run **SVM.POT**-leave
'We will run away.' (Elicitation_100719_AB)

(13) ***máà**-se* ***ɛ**-sɛ̀*
**1s.POT**-run **SVM.POT**-leave
'I will run away.' (Elicitation_100714_QM)

In other aspects and moods, the form of the serial verb marker is é-/ɛ́-; for instance, in the progressive in (14a) and (15), and the habitual in (14b) and (16).

(14) a. ***wɛ̀ɛ́**-gà* ***ɛ́**-za*
**2s.PROG**-walk **SVM.PROG**-pass
'You are passing through.'
b. ***wụ-zɛ̌**-gà* ***ɛ́**-za*
**2s.HAB-HAB**-walk **SVM.HAB**-pass
'You pass through (all the time).'

(15) ***bɛ̀ɛ́**-ŋwya* ***ɛ́**-kpɛ*
**$C_1$p.PROG**-throw **SVM.PROG**-put.in
'They were throwing (it) in.' (Folkstory_110406_QM_01:29)

(16) *ńte* *mè* *sị̀* *ba* *petee* ***bị́-zɛ̌**-za*
LOC:like.that inside COMP $C_1$p.SBJ all **$C_1$p.HAB-HAB**-pass
***ɛ́**-klanị̀* ***ɛ́**-pɛ* *bị-ŋà∼ŋà=wɛ*
**SVM.HAB**-go.around **SVM.HAB**-want $C_4$p-REDUP∼eat=DEF
*ní* *li-ŋwàfụ=nɛ* *mɛ̀*
LOC $C_3$s-forest=DEF inside
'So they all used to roam around the forest looking for food.'
(Folkstory_110406_QM_00:21)

Example (17) shows é-/ɛ́- is also used in the negative.

(17) *kù-ni = o* ***kú-lí****-kpɛ* ***é****-ple*
$C_5$S-water = DEF **$C_5$S.PROG.NEG-PROG.NEG**-put.in **SVM.PROG**-descend
'The water doesn't flow (in the river).' (Avatime-history_110905_BB_07:20)

The optional recurrent aspect and intentive mood modifiers do not influence the serial verb markers. For instance, the serial verb markers in examples (18) and (19) are truncated forms of the subject agreement prefixes as would typically be found in the perfective, rather than the ɔ- or ɛ́- forms found with the other contrastive aspects and moods.

(18) *yɛ́* *sị̀* ***bɛ-tá****-kɔ̀* ***ɛ****-wà* *kunu = yè*
$C_1$S.SBJ:FOC COMP **$C_1$p.PFV-INT**-take **SVM.$C_1$p.PFV**-use funeral = DEF
'He is the one they will use for the funeral.' (Folkstory_110406_QM_06:29)

(19) *a-zɛ̌-se* *a-trɛ* *ní* *ɔ̀vanɔ̀*
**$C_1$S.PFV-REC**-run **SVM.$C_1$S.PFV**-go LOC Vane
'He was running to Vane.'

As mentioned at the beginning of this section, the use of serial verb markers is optional and has no apparent semantic or pragmatic influence. The frequency of their use varies according to age. Older speakers use them very frequently, and younger speakers use them only rarely, if at all. This suggests they may be lost in future generations and that the apparent optionality is linked with this shift in the language rather than having an independent functional motivation.

### 3.2.2 Aspect and mood within SVCs

Typically, all verbs in an SVC must share a single value for aspect and mood (Aikhenvald 2006; Durie 1997). This restriction is, however, reported to be relaxed among Kwa languages where it is common for each verb phrase to be individually modifiable for aspect and modality (Ameka 2003). In particular, the local lingua franca Ewe (Ameka 2006a) and Avatime's close relative Tafi (Bobuafor 2013:302) both allow independent marking of aspect and modality on each verb within an SVC as long as it is semantically plausible. One could, thus, expect Avatime to follow the local pattern and allow independent marking of aspect and modality within SVCs. However, that is not the case. All Avatime clauses, monoverbal and SVC, must be

marked for one and only one of the six contrastive aspects and moods: perfective, progressive, habitual, potential, subjunctive, and imperative (Defina in press).

The markers for all six contrastive aspects and moods have historically fused with the subject agreement prefixes so that both categories are now indicated using the one agreement prefix as described in Section 2.7.2.1. SVCs in each of the aspect and mood categories are shown below in examples (20)-(25).

(20) *bɛ-sị́ níklɔ ɛ-pɛ̀ kị́=ba*
$C_1$p.PFV-say there $C_2$s.PFV-be.good give=$C_1$p.OBJ
'They said as for that place, it was good for them.'
(Avatime-history_110905_BB_11:24)

(21) *kìí-dzi trɛ kè-de=a tàe*
1p.PROG-return go $C_6$s-back=DEF a.little
'We are going back a little.' (Midwifery_110901_AB_07:13)

(22) *bá-nɔ̀=atɔ be-zè gì ba kóŋ*
$C_1$p-person=INDEF $C_1$p.PFV-be.NPRES REL $C_1$p.SBJ at.all
***bị́-zɛ̌-pɔ kị́ bá-nɔ̀=a***
**$C_1$p.HAB-HAB-help give $C_1$p-person=DEF**
'There were special people who used to help people (deliver babies)'
(Midwifery_110901_AB_01:11)

(23) *bíà-kɔ manị̀ be-bi=wà*
$C_1$p.POT-take bring $C_1$p.POS-child=DEF
'They will bring (it) to their children.' (Folkstory_110406_QM_00:25)

(24) *kɔ bí-zizi wɔ ke-le=a mè petee kị́=wɔ=ɛ*
so $C_1$p.SUBJ-spoil 2s.POS $C_6$s-world=DEF inside all give=2s.OBJ=CM
'So they'll spoil all your life for you.' (Midwifery_110901_AB_07:40)

(25) *kɔ bɛ-sị́ kpɛ ple-nò*
so $C_1$p.PFV-say [IMP]put.in descend-COM
'So they said "push down!".' (Midwifery_110901_AB_01:58)

The habitual is the only form which has a separate marker (example (22)), but even here the *zě-/zɛ̌-* prefix is not sufficient for indicating the aspect since it is homophonous with the recurrent and itive prefixes and thus requires the agreement prefix for disambiguation. In the case of the imperative (example (25)), the mood is marked by the absence of the agreement prefix. This fusion of aspect and mood marking with subject agreement makes it impossible to use the standard strategies for marking aspect and mood in monoverbal clauses with subsequent verbs in SVCs.

It is also not considered grammatical to use a serial verb marker to indicate an aspect or mood different to that of the first verb, as can be seen in example (26). Example (a) attempts to combine a perfective marked first verb with a potential marked serial verb marker. Example (b) attempts to combine a perfective marked first verb with the *ɛ́-* serial verb marker which can indicate progressive, habitual, or subjunctive. Habitual would not be semantically plausible in this case, but progressive and subjunctive should be with an interpretation such as 'You left Gbadzeme and are coming to Vane' or 'You left Gbadzeme to come to Vane' respectively. Example (c) attempts to combine a potential marked first verb with a perfective serial verb marker on the second verb. All three sentences are regarded as ungrammatical.

| | | | | | |
|---|---|---|---|---|---|
| (26) | a. | **mlɛ-dɔ* | *Gbàdzɛmɛ̀* | *ɔ-ba* | Ɔ̀vanɔ̀ |
| | | 2p.PFV-move.from | Gbadzeme | SVM.POT-come | Vane |
| | b. | **mlɛ-dɔ* | *Gbàdzɛmɛ̀* | *ɛ́-ba* | Ɔ̀vanɔ̀ |
| | | 2p.PFV-move.from | Gbadzeme | SVM.PROG/SUBJ-come | Vane |
| | c. | **kɪ́à-dɔ* | *Gbàdzɛmɛ̀* | *ɪ̀-ba* | Ɔ̀vanɔ̀ |
| | | 1p.POT-move.from | Gbadzeme | SVM.1p.PFV-come | Vane |

(Elicitation_100717_AB)

It is thus not possible to independently mark verbs within Avatime SVCs for aspect or mood categories from this contrastive set. However, Avatime has two optional categories: the recurrent and intentive. These can be additionally marked on any simple monoclausal verb and since they are marked by independent morphemes, they do not face the same practical limitations.

The recurrent aspect, used for indicating repeated action, can in fact modify each part of an SVC independently. For instance, the (a) examples in (27) and (28) have the recurrent marked on the first verb *dzɛ* 'go' and the going is necessarily repeated,

while the (b) examples have the recurrent marked on the second verb *wà* 'work' and it is only the working and not the going which is repeated.

(27) a. *mà-**zɛ̌**-dzɛ* *Òholò* *a-wà* *à-xwɛ̀=na*
1s.PFV-**REC**-go Ho SVM.1s.PFV-work $C_3$p-job=DEF
'I was going to Ho and working.' (went and returned repeatedly)

b. *mà-dzɛ* *Òholò* *a-**zɛ̌**-wà* *à-xwɛ̀=na*
1s.PFV-go Ho SVM.1s.PFV-**REC**-work $C_3$p-job=DEF
'I went to Ho and was working.' (moved there for some time)

(28) a. *máà-**zɛ̌**-dzɛ* *Òholò* *ɔ-wà* *à-xwɛ̀=na*
1s.POT-**REC**-go Ho SVM.POT-work $C_3$p-job=DEF
'I will be going to Ho and working.' (coming and going repeatedly)

b. *máà-dzɛ* *Òholò* *ɔ-**zɛ̌**-wà* *à-xwɛ̀=na*
1s.POT-go Ho SVM.POT-**REC**-work $C_3$p-job=DEF
'I will go to Ho and be working.' (move there for some time)

This is only possible with some SVCs, such as those expressing sequential action (See Section 3.4.3). For example, according to speaker reports, the SVC in (29) must describe separate consecutive actions of making and giving rather than the benefactive interpretation of making rice for the people.

(29) *mà-tɔ́* *kɪ̣-mɪ̣mɪ̣* *kpáùŋ* *a-**zɛ̌**-kɪ́* *bá-nɔ̀=a*
1s.PFV-cook $C_4$s-rice plenty SVM.1s.PFV-**REC**-give $C_1$p-person=DEF
'I made plenty of rice and was giving it to the people.'

In the case of the intentive mood, however, there appears to be a more fundamental restriction. The intentive marker *tá-* can only occur on the first verb of an SVC and it must scope over the whole construction, as shown in the following examples. Example (30a) shows an SVC with the first verb marked with the intentive. Discussions with consultants showed the intentive does not narrowly apply to the first verb *kɔ̀* 'take' since the sentence can also be used when the speaker has picked up the axe but not (yet) used it for splitting the firewood. Example (30b)

shows it is not possible to specify this interpretation by marking the intentive on the second verb. This can only be done using two separate clauses as in (30c).

(30) a. *ma-**tá**-kɔ̀* *kà-wɛ = a* *a-yài* *ɔ̀-nyị = nɔ̀*
1s.PFV-**INT**-take $C_6$s-axe = DEF SVM.1s.PFV-break $C_2$s-firewood = DEF
'I intend to split the firewood with the axe'/'I intend to take the axe and split the firewood.' (can be used regardless of whether or not the speaker is already holding the axe)

b. **ma-kɔ̀* *kà-wɛ = a* *a-**tá**-yài* *ɔ̀-nyị = nɔ̀*
1s.PFV-take $C_6$s-axe = DEF SVM.1s.PFV-**INT**-break $C_2$s-firewood = DEF
Intended: 'I took the axe and intend(ed) to split the firewood.' or 'I intend(ed) to split the firewood with the axe (which I am already holding)'.

c. *ma-kɔ̀* *kà-wɛ = a.* ***ma-tá**-yài* *ɔ̀-nyị = nɔ̀*
1s.PFV-take $C_6$s-axe = DEF **1s.PFV-INT**-break $C_2$s-firewood = DEF
'I took the axe. I intend(ed) to split the firewood.'

This restriction even holds with SVCs more clearly referring to sequential actions, as demonstrated by example (31). The sentence in (a) shows a sequential action SVC with the intentive marked on the first verb and scoping over both actions. Example (b) shows the intentive cannot be marked on the second verb to give a narrow scope reading. This can again only be achieved by using separate clauses as in (c and d).

(31) a. *mà-**tá**-dzɛ* *Òholò* *a-wà* *à-xwɛ̀ = na*
1s.PFV-**INT**-go Ho SVM.1s.PFV-work $C_3$p-job = DEF
'I intend to go to Ho and work.'

b. **mà-dzɛ* *Òholò* *a-**tá**-wà* *à-xwɛ̀ = na*
1s.PFV-go Ho SVM.1s.PFV-**INT**-work $C_3$p-job = DEF
Intended: 'I went to Ho intending to work.'

c. *mà-dzɛ* *Òholò* ***ma-tá**-wà* *à-xwɛ̀ = na*
1s.PFV-go Ho **1s.PFV-INT**-work $C_3$p-job = DEF
'I went to Ho and intend to work.'

d. *mà-dzɛ* *Òholò* ***tɔ*** ***máà**-wà* *à-xwɛ̀ = na*
1s.PFV-go Ho **PURP** **1s.POT**-work $C_3$p-job = DEF
'I went to Ho in order to work.'

### 3.2.3 Argument sharing within SVCs

The verbs in Avatime SVCs share a single grammatical subject. This subject must also be the actor argument for all verbs in the SVC, as in (32).

(32) *lɛ̌* *be-dzì* *mu=i*
then $C_1$p.PFV-return ascend=CM
'Then they ascended again.' (Avatime-history_110905_BB_01:59)

This means Avatime does not allow the sort of switch subject resultative SVCs commonly found in many other serializing languages, where the second verb is intransitive and takes the undergoer of the first verb as its sole argument (33). These sorts of resultative meanings may, however, be expressed using SVCs in Avatime if the second verb is transitive, as in (34), or labile, as in (35), so the subject of the first verb is the actor for both verbs. A similar restriction is also reported in Ewe and Likpe (Ameka 2003).

(33) **a-ta* *ɔ-ga=ɛ* *tse*
$C_1$s.PFV-shoot $C_1$s-goat=DEF die
Intended: 'He shot the goat dead.'

(34) *a-ta* *ɔ-ga=ɛ* *ye*
$C_1$s.PFV-shoot $C_1$s-goat=DEF kill
'He shot the goat dead.'

(35) a. *wò-trutru* *ò-pupo=lò* *dra*
2s.PFV-push $C_2$s-door=DEF open
'You pushed the door open.'
b. *wɔ̀-dra* *ò-pupo=lò*
2s.PFV-open $C_2$s-door=DEF
'You opened the door.'
c. *ò-pupo=lò* *ɛ̀-dra*
$C_2$s-door=DEF $C_2$s.PFV-open
'The door is open.'

Other arguments may also be shared by verbs within SVCs. In these cases, they are mentioned only with their first verb. For instance, in example (36) the object of the first verb (*lìfiflìnɛ* 'a type of porridge') is shared by the subsequent three verbs.

(36) *xé* *bɛ̀ɛ́-bɔ* ***lì-fiflì=nɛ***
when $C_1$p.PROG-mould **$C_3$s-type.of.porridge=DEF**
*ɛ́-nywà* *ɛ́-kpɛ* *ɛ́-kí* *ɔ-kà-tsì=e*
SVM.PROG-throw SVM.PROG-put SVM.PROG-give $C_1$s-father-old=DEF
'When they were moulding the porridge and threw it to the old man.'
(Folkstory_110406_QM_03:25)

### 3.2.4 Focus and SVCs

As is common for West African SVCs (Ameka 2003), individual verbs in Avatime SVCs can be marked for focus. Most commonly, focus is marked on the first verb. This can have narrow scope over that particular verb, as in (37). Here speaker A asks a question consisting of two clauses, the first has focus on the verb *gà* 'walk' the second clause contains an SVC with 'run' and 'go'. Speaker B responds with a single SVC clause with focus on the first verb.

(37) *A:* *ki-gá* *afua* *a-gà* *aló* *e-se* *trɛ*
VFOC-walk:FOC Afua $C_1$s.PFV-walk or $C_1$s.PFV-run go
*ní* *kè-dzi=à* *mɛ̀* *na?*
LOC $C_6$s-market=DEF inside QM
'Did Afua walk or run to the market?'
*B:* ***ki-sé*** ***e-se*** *trɛ*
**VFOC-run:FOC** **$C_1$s.PFV-run** go
'She [ran]$_{FOC}$ to the market.'

Alternatively, it can have broad scope over the whole SVC or even relate more to the second verb, as in (38), where the important point is really the leaving rather than the getting up.

(38) *i-mɔ̀* *àsafò* *ye-bi=à* ***ki-yɔ́*** ***bɛ-yɔ́*** ***sɛ́*** *lo*
?-see Asafo $C_1$s.POS-child=DEF **VFOC-get.up:FOC** **$C_1$p.PFV-get.up** **leave** FP
'Look at Asafo's children, they [got up and left]$_{FOC}$.' (Conv-street_100720-2)

It is also possible for non-initial verbs to be focused, as in (39). This appears to be much less common and so far is only attested in elicitation sessions after prompting.

| (39) | ***ki-nú*** | *ɔ-nṵ̀vɔ̀=yɛ* | *o-dí* | ***nu*** | *bɔl=yɛ* |
|---|---|---|---|---|---|
| | **VFOC-listen:FOC** | $C_1$S-child=DEF | $C_1$S.PFV-sit | **listen** | football=DEF |

'The boy sat [listening]$_{FOC}$ to the football.'

### 3.2.5 Summary of Avatime SVC properties

Avatime SVCs are characterized by the following properties:

i) A sequence of two or more verbs in a single clause
ii) No predicate-argument relation between the verbs
iii) Only the first verb is fully inflected for subject agreement, aspect, mood, and polarity
iv) Subsequent verbs may be bare or prefixed with a reduced agreement marker
v) The recurrent aspect may be independently marked on subsequent verbs in some SVCs, all other aspects and moods may be marked only once on the first verb and scope over the whole construction.
vi) The subject must be an argument of all verbs
vii) Other arguments may also be shared by verbs within the SVC, in which case they are mentioned once only following their first verb.
viii) Individual verbs may be focused

Avatime SVCs thus exhibit many properties typical of West African serial constructions. In particular, their ability to focus individual verbs, the lack of switch subject resultative type SVCs, and the ability to independently mark subsequent verbs for aspect. They also differ from the typical West African pattern in their use of reduced agreement markers on subsequent verbs and the more generally typical restrictions on independent aspect and mood marking. In the next section, I examine the different functions for which SVCs are utilized in Avatime.

## 3.3 Semantic functions of Avatime SVCs

Avatime SVCs are employed for a broad range of semantic functions. These functions can be divided into three broad groups: modifying, argument adding or

marking, and grouping consecutive actions. These semantic groupings also have subtly different syntactic properties as discussed further in Section 3.4.

In modifying SVCs, the first verb modifies the way the action described by the second verb is performed. These include the typical manner plus path of motion SVCs, as in (40) where the first verb describes the manner and the second the path.

(40) *kò* *e-se* *dɔ* *nílɔ* *gì* *e-kpò=e*
then $C_1$S.PFV-move.quickly move.from there REL $C_1$S.PFV-hide=CM
'Then he ran out from where he was hiding.' (Folkstory_110406_QM_03:48)

Another kind of modifying SVC is where the first verb describes the posture of the actor during the action or state described by the second verb, as in (41).

(41) *o-di* *ŋwɛ̀*
$C_1$S.PFV-sit drink
'S/he sits drinking.'

I also include among the modification SVCs some slightly less typical cases, such as when two path verbs combine in an SVC to create a complex path, as in (42) and (43). Note in these cases the two path elements combine simultaneously with the first verb modifying the way the motion described by the second verb is carried out. They thus fall within the modifying SVC category rather than the sequential category discussed below.

(42) *lɛ̌* *ba-tɔ* *be-ple* *e-ku* *ní* *Gbàdzɛmɛ̀=ɛ*
and $C_1$p-INDEF $C_1$p.PFV-descend SVM.$C_1$p.PFV-enter LOC Gbadzeme=CM
'And some descended into Gbadzeme.' (Avatime-history_110905_BB_10:42)

(43) *lɛ̌* *be-dzì* *mu=i*
and $C_1$p.PFV-return ascend=CM
'Then they ascended again.' (Avatime-history_110905_BB_01:59)

Finally, I also include SVCs such as (35), repeated here as (44), and (45) within the modification type of SVCs. In these SVCs, the first verb is an activity verb describing the manner of achieving the action described by the second verb. They are often

translated using a resultative construction. However, since the second verb is transitive, they have more in common with the manner plus path SVCs than with resultative SVCs in many other languages where the result is generally expressed by an intransitive stative verb.

(44) *wo-trutru* *ò-pupo = lò* *dra*
2s.PFV-push $C_2$s-door = DEF open
'You pushed the door open.'

(45) *bèɛ́-ŋwya* *ɛ́-kpɛ*
$C_1$p.PROG-throw SVM.PROG-put.in
'They were throwing (it) in.' (Folkstory_110406_QM_01:29)

The second major type of semantic function for Avatime SVCs is argument adding or marking. There are two types of argument adding SVCs. In both, the choice of verb is fixed and the construction is moving towards grammaticalization. One type is formed with the 'give' verb *kị* as the second verb and is used to add a benefactor or recipient role, as in (46). The other type uses the 'take' verb *kɔ̀* as the first verb and adds an instrument, means, or manner, as in (47)-(49) respectively.

(46) *bɛ-plɛ* *ɔ̀-tɔsì = lɔ* *ɛ-kí*[1] *ò-kusì = e*
$C_1$p.PFV-put.down $C_2$s-bed.mat = DEF SVM.$C_1$p.PFV-give $C_1$s-chief = DEF
'They laid a bed for the chief.' (Avatime-history_110905_BB_09:29)

(47) *a-kɔ̀* *kà-wɛ = a* *yài* *ò-se = lò*
$C_1$s.PFV-take $C_6$s-axe = DEF break $C_2$s-tree = DEF
'He used an axe to split the tree.'

[1] Note the tone on the 'give' verb *kị* is raised from high to extra-high. This is done whenever 'give' is used as a non-initial verb in an SVC regardless of whether it has a benefactive function, as is the case here, or its full literal meaning as in example (50)

(48) *a-kɔ̀* *ku-zò* *dzi* *ohonete*
$C_1$s.PFV-take $C_5$s-theft become rich.person
'Through theft he became a rich man.'

(49) *a-kɔ̀* *ku-siyeyome* *sɛ̀*
$C_1$s.PFV-take $C_5$s-anger leave
'He left in anger.'

In addition to adding a new argument, 'take' verbs can be used to mark the theme in a three-place predicate, as in (50). Note the theme argument is already an argument of the second verb; it is not added by the take verb, as shown in (50b). Such object marking is a very common function of take-SVCs crosslinguistically, both within and outside West Africa (Lord 1993). Note these take-SVCs are the focus of the studies in Chapter 6.

(50) a. *a-kɔ̀* *lị-ba=lɛ* *kí=yɛ*
$C_1$s.PFV-take $C_3$s-hoe=DEF give=$C_1$s.OBJ
'He gave him the hoe.' (Folkstory_110406_QM_05:24)

b. *a-kị=yɛ* *lị-ba=lɛ*
$C_1$s.PFV-give=$C_1$s.OBJ $C_3$s-hoe=DEF
'He gave him the hoe.'

The 'take' verbs used to introduce new arguments are more semantically bleached and grammaticalized than the 'take' verbs used to mark themes in three-place predicates. When introducing new arguments, only the generic *kɔ̀* 'take' verb can be used. In contrast, when marking a theme any of the verbs of taking can be used depending on the kind of object and how it was taken, as can be seen in (51) where the 'gather' verb is used with the theme rice.

(51) *kɔ* *bɛ-halị̀* *a-mụ=nà* *ɛ-manɔ̀*
so $C_1$p.PFV-gather $C_3$p-rice=DEF SVM.$C_1$p.PFV-bring
*ní* *ke-tsripà* *mɛ̀* *ní* *ɔ̀-nyɔ=nɔ̀* *mɛ̀*
LOC $C_6$s-clearing inside LOC $C_2$s-farm=DEF inside
'They brought the rice to the clearing in the farm.'
(Rice-farming_100613_EN_05:26)

The final function of Avatime SVCs is to combine consecutive actions, as in (52). These sequential actions must form a single culturally relevant and cohesive unit, generally with an overarching goal. For instance, the source and goal of a motion event, as in (53), or the actions required for achieving a task such as cooking a meal, as in (54). When there is no such overarching goal, the actions can only be combined using coordinated sentences, as in (55).

(52) *lɛ̌* *a-ya=lɛ* *e-dù=i*
then $C_1$S.PFV-divide=$C_3$S.OBJ SVM.$C_1$S.PFV-put=CM
'Then she divided it (the porridge) and put it down.'
(Folkstory_110406_QM_01:14)

(53) *ɔ-dzɛ* *a-dɔ* *ɔ̀-ma=nɔ̀* *mɛ̀* *ba* *sku*
$C_1$S-woman $C_1$S.PFV-exit $C_2$S-town=DEF inside come school
'The woman left the town and came to school.'

(54) *ma-tsà* *tomatoes=ye* *a-kpɛ*
1S.PFV-cut tomatoes=DEF SVM.1S.PFV-put
*ní* *kè-zi=a* *mɛ̀*
LOC $C_6$S-bowl=DEF inside
'I cut tomatoes and put them in the bowl.'

(55) a. **ma-tsà* *tomatoes=ye* *a-sɛ́*
1S.PFV-cut tomatoes=DEF SVM.1S.PFV-leave:LOC
*ke-pe=a* *mɛ̀*
$C_6$S-house=DEF inside
Intended: 'I cut tomatoes and left the house.'
b. *ma-tsà* *tomatoes=ye*
1S.PFV-cut tomatoes=DEF
***lɛ̌*** *ma-sɛ́* *ke-pe=a* *mɛ̀*
**and** 1S.PFV-leave:LOC $C_6$S-house=DEF inside
'I cut tomatoes and left the house.'

This restriction is a common property of sequential action SVCs and has been noted by several researchers for other languages (e.g. Bruce 1988; Diller 2006; Durie 1997; Enfield 2002; Jarkey 1991; Lewis 1993). Note this type of SVC features extensively in Chapter 5.

## 3.4 Subtypes of SVCs in Avatime

The SVCs used for the different types of functions – modifying, argument marking and adding, and combining sequential actions – have subtly different morpho-syntactic properties, see Table 3.2. These differences divide Avatime SVCs into three subtypes – nuclear, core, and sequential – which to a large extent, but not exactly, mirror the functional divisions, see Table 3.3. I discuss each of the subtypes in detail below.

Table 3.2: Morphosyntactic properties distinguishing the subtypes of Avatime SVCs

| **Property** | | **Nuclear** | **Core** | **Sequential** |
|---|---|---|---|---|
| Can aspectual adverbials occur | between verbs? | Marginally | Yes | Yes |
| | with restricted scope? | No | Yes | Yes |
| Can directionals occur | on subsequent verbs? | Marginally | Yes | Yes |
| | with restricted scope? | No | Not clear | Yes |
| Can the recurrent occur | on subsequent verbs? | No | Yes | Yes |
| | with restricted scope? | No | Not clear | Yes |
| Can locational or temporal adverbials occur | between verbs? | No | Yes | Yes |
| | with restricted scope? | No | No | Yes |

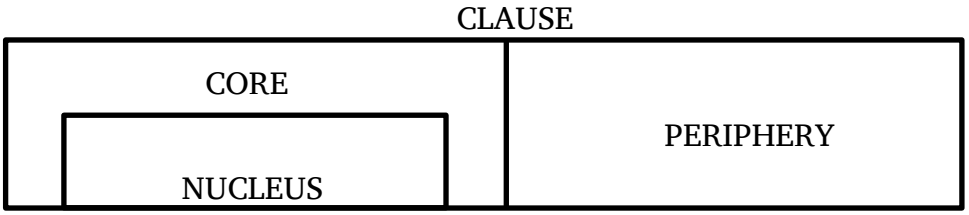
Table 3.3: Semantic functions of SVCs in each subtype

| **Subtype** | **Semantic functions** |
|---|---|
| Nuclear | Modifying (Posture, Manner + path, Complex path, and Manner + activity)<br>Theme marking |
| Core | Argument adding and theme marking<br>Modifying (Manner + activity) |
| Sequential | Combining consecutive actions |

Since the Role and Reference Grammar (RRG) analysis of SVCs (Foley & Olson 1985; Foley & Van Valin 1984; Van Valin 2005) closely matches the observed Avatime patterns, it is used as a helpful way of describing the data. I shall, thus, briefly introduce the essential aspects of RRG before continuing with the description of SVC subtypes in Avatime. There are two main ideas behind the RRG analysis of complex clauses: nexus types and the layered structure of the clause (Van Valin & Foley 1980; Van Valin 2005). There are three nexus types, or ways of joining elements together. These are the standard subordination and coordination, and an additional type called cosubordination, which combines coordination like properties with the operator dependence typical of subordination (Van Valin 2005:187)[2]. RRG also divides clauses into a layered structure of three parts (see Figure 3.1). The nucleus is the bare predicate and in a simple clause would consist of the verb stem only. The core consists of the predicate plus any core arguments. The periphery contains any non-core arguments, such as adjuncts. The nucleus is contained within the core, the periphery is adjoined to the core, and all together they make up the clause. These nexus types and parts of the clause interact to form different kinds of complex clauses, for instance two nuclear level units can be joined via subordination or coordination, or they could connect two periphery level units.

[2] Though there have been some arguments against cosubordination as a distinct nexus type (Bickel 2010; Foley 2010b)

Figure 3.1: The RRG layered structure of the clause




(Van Valin 2005:4)

In nuclear SVCs, the verbs are joined together, via cosubordination, under a single nucleus node. This is a tight bond and the verbs in these SVCs cannot be modified individually by nuclear level operators, such as aspectual modifiers. Core SVCs, in contrast, contain two separate nuclear nodes each within separate core nodes which are joined, via cosubordination, under another core node. The verbs in these SVCs can have their own distinct core arguments and be individually modified by nuclear, but not core, level operators (Foley & Olson 1985).

### 3.4.1 Nuclear SVCs in Avatime

The core members of this group are SVCs where the first verb modifies the way the action described by the subsequent verb is carried out i.e. manner, posture, or complex path. Additionally, SVCs where a 'take' verb marks the theme argument of a three-place predicate sometimes behave like these modifying SVCs though they may also behave like the argument adding SVCs discussed in the next section. This variation does not mean they have some traits of nuclear and some traits of core SVCs. Rather a speaker sometimes responds to all questions consistently as if the construction behaves like the modifying SVCs and other times responds as if it behaves in the same way as the argument adding SVCs.

Nuclear SVCs are the most restrictive, with the tightest connection between their verbs. The verbs cannot be independently modified and adverbials do not typically appear between them. Directional affixes can occur on one verb only, typically the first, and always scope over the whole construction, as can be seen in (56) and (57)[3].

[3] Asterisks indicate that all informants consistently rejected the sentence. Question marks indicate that most consultants rejected the sentence as ungrammatical but at least one speaker at one time accepted it. Other sentences were accepted by all informants. All sentences were tested with between three and five consultants.

In these two examples, the SVCs in (a) have a first verb marked with a directional which scopes over the whole construction. The SVCs in (b) show the dispreference for placing a directional on the second verb: the itive in (56) was rejected by all consultants, while the ventive in (57) was rejected by two out of three consultants.

(56) a. *Komla* *a-**ʒɛ**-tà* *ɔ-gà=ɛ* *ye*
Komla $C_1$s.PFV-**IT**-shoot $C_1$s-goat=DEF kill
'Komla went and killed the goat.'

b. **Komla* *a-tà* *ɔ-gà=ɛ* ***ze**-ye*
Komla $C_1$s.PFV-shoot $C_1$s-goat=DEF **IT**-kill
Intended: 'Komla went and killed the goat.'

(57) a. *Komla* *a-**bá**-kɔ̀* *kù̩-sà* *kí̩* *ɔ́-dzɛ*
Komla $C_1$s.PFV-**VENT**-take $C_5$s-cloth give $C_1$s-woman
'Komla came and gave the cloth to the woman.'

b. ?*Komla* *a-kɔ̀* *kù̩-sà* ***bá**-kí̩* *ɔ́-dzɛ*
Komla $C_1$s.PFV-take $C_5$s-cloth **VENT**-give $C_1$s-woman
'Komla came and gave the cloth to the woman.'

The recurrent aspect – the only aspect that can independently modify verbs within some Avatime SVCs (Section 3.2.2) – can only be marked on the first verb in nuclear SVCs, as in (58).

(58) a. *ba-**ʒɛ̌**-dí̩* *ŋwɛ̀* *kù-gòda*
$C_1$p.PFV-**REC**-sit drink $C_6$s-palmwine
'They were sitting drinking palm wine.'

b. **ba-dí̩* ***ʒɛ̌**-ŋwɛ̀* *kù-gòda*
$C_1$p.PFV-sit **REC**-drink $C_6$s-palmwine

Aspectual adverbials between the verbs are strongly dispreferred according to speaker judgements, and regardless of position always scope over the entire construction, as can be seen in (59).

(59) a. ***koko*** *ba-dí* *gu* *ku-nugu=yò*
**already** $C_1$p.PFV-sit talk $C_5$s-mouth=DEF
'They already sat talking.'

b. *ba-dí* *gu* *ku-nugu=yò* ***koko***
$C_1$p.PFV-sit talk $C_5$s-mouth=DEF **already**
'They already sat talking.'

c. ?*ba-dí* ***koko*** *gu* *ku-nugu=yò*
$C_1$p.PFV-sit **already** talk $C_5$s-mouth=DEF
'They already sat talking.'

While aspectual adverbials are dispreferred between the verbs, they are occasionally accepted. as in (59c). In contrast, locative and temporal adverbials are never accepted between verbs, for instance (60) and (61). In all cases the adverbial scopes over the whole SVC.

(60) a. *me-feke* *lị-kla=nɛ̀* *vù* ***níyà***
1s.PFV-pick.up $C_3$s-stone=DEF hold **here**
'I picked up the stone here.'

b. **me-feke* *lị-kla=nɛ̀* ***níyà*** *vù*
1s.PFV-pick.up $C_3$s-stone=DEF **here** hold
Intended: 'I picked up the stone here held it.'

(61) a. *ba-dí* *gu* *ku-nugu=yò* ***kivòe***
$C_1$p.PFV-sit talk $C_5$s-mouth=DEF **yesterday**
'They sat talking yesterday.'

b. **ba-dị* ***kivòe*** *gu* *ku-nugu=yò*
$C_1$p.PFV-sit **yesterday** talk $C_5$s-mouth=DEF
Intended: 'They sat yesterday talking.'

The fact that aspect and directionals must scope over both verb phrases suggests the verbs are within a single nucleus in RRG terms (Van Valin 2005). In the early days of RRG, it was believed nuclear SVCs would also require the verbs to be adjacent with no intervening object noun phrase (Foley & Olson 1985). However, several languages have since been documented where nuclear SVCs allow object

NPs to occur between the verbs (Bril 2004; Crowley 2002; Durie 1997). Avatime is the first example of this within West Africa, a region which was previously claimed to lack nuclear SVCs due to the believed adjacency restriction (Foley & Olson 1985).

### 3.4.2 Core SVCs in Avatime

The principle members of this group are argument-adding SVCs. Additionally, SVCs with a 'take' verb used to mark a theme argument and the manner plus action type of modifying SVCs can also occur in this group. Avatime core SVCs typically allow subsequent verbs to be independently modified. Adverbials can occur between the verbs, though locational and temporal adverbials scope over the whole construction.

Subsequent verbs in core SVCs can be marked with directionals. This is shown in example (62) where consultants reported a difference in meaning between the sentences in (a) and (b). Placing the directional on the second verb in (b) gives the construction more of a sequential action rather than a purely instrumental reading. Given the semantics of the directional, a narrow scope reading would automatically lead to a sequential action interpretation. It is not clear whether a narrow scope use of the directional shifts the SVC over to the sequential subtype (Section 3.4.3). Further testing with locational and temporal adverbials would be needed to discriminate these two possibilities. It is, however, notable that narrow scope use of the directionals is not possible with nuclear SVCs even though such a shift in interpretation would be semantically plausible.

(62) a. *ma-**ʒɛ**-kɔ̀* *kà-wɛ=a* *tsà* *ò-se=lò*
1s.PFV-**IT**-take $C_6$s-axe=DEF cut $C_2$s-tree=DEF
'I went and used the axe to cut the tree.'

b. *ma-kɔ̀* *kà-wɛ=a* ***ʒɛ**-tsà* *ò-se=lò*
1s.PFV-take $C_6$s-axe=DEF **IT**-cut $C_2$s-tree=DEF
'I took the axe to go cut the tree.'

The recurrent aspect can be placed on subsequent verbs in core SVCs. When it is used with benefactive SVCs, a sequential action reading is given, as in (63). When it is used with instrumentative SVCs, the recurrent can have narrow scope without leading to a sequential action meaning, as in (64). In (64b), it is the time of cutting the tree that is in focus with repeated cutting actions on this tree. The axe was used

for at least some of this time, but may have only been one of the tools used. Whereas in (64a), it is the using of the axe which is in focus and repeated. It may have also been used to cut other trees or for other purposes.

(63) *mà-tɔ* *kị̀-mịmị̀* *kpáùŋ* ***zɛ̌-kí*** *bá-nɔ̀ = a*
1s.PFV-cook $C_4$s-rice plenty **REC**-give $C_1$p-person = DEF
'I made plenty of rice and was giving it to people.'

(64) a. *a-**zɛ̌**-kɔ̀* *kà-wɛ = a* *tsà* *ò-se = lò*
$C_1$s.PFV-**REC**-take $C_6$s-axe = DEF cut $C_2$s-tree = DEF
'He was using the axe to cut the tree.'
b. *a-kɔ̀* *kà-wɛ = a* ***zɛ̌-tsà*** *ò-se = lò*
$C_1$s.PFV-take $C_6$s-axe = DEF **REC**-cut $C_2$s-tree = DEF
'He used the axe and was cutting the tree.'

Aspectual adverbials can easily occur between verbs in these SVCs and can have restricted scope over one verb phrase only, as can be seen in (65). Here the adverb *koko* 'already' modifies the first verb phrase *ebu àgbèlìye ìdrulè* 'he dug cassava mounds', but does not scope over the second verb phrase *kí Kwami* 'give to Kwami' which may be yet to occur.

(65) *e-bu* *àgbèlì = ye* *ì-dru = lè* ***kóko*** *kí* *Kwami*
$C_1$s.PFV-remove cassava = DEF $C_2$p-mound = DEF **already** give Kwami
'He already dug cassava mounds for Kwami.' (I.e. the cassava mounds have been dug, possibly without the intention of giving them over to Kwami.')

Locative and temporal adverbials can also occur between verbs in core SVCs. Unlike aspectual adverbials, they must scope over the entire SVC unless they modify one of the nominal arguments. So the pairs of SVCs in (66) and (67) have the same meaning, though the locational adverbial in (67b) can also be interpreted as modifying the nominal 'cloth' rather than the action.

(66) a. *mà-dzɛ* *Òholò* *kí* *Akosua* ***kivòe***
1s.PFV-go Ho give Akosua **yesterday**
'I went to Ho for Akosua yesterday.'

*b.* *mà-dzɛ* *Òholò* ***kivòe*** *kį́* *Akosua*
1s.PFV-go Ho **yesterday** give Akosua
'I went to Ho for Akosua yesterday.'

(67) a. *a-kɔ̀* *kù̱-sà = a* *kį́* *ɔ́-dzɛ = ɛ*
$C_1$s.PFV-take $C_5$s-cloth = DEF give $C_1$s-woman = DEF
***ní*** ***ke-pe = a*** ***mɛ̀***
**LOC** **$C_6$s-house = DEF** **inside**
'He gave the cloth to the woman in the house.'

b. *a-kɔ̀* *kù̱-sà = a* ***ní*** ***ke-pe = a*** ***mɛ̀***
$C_1$s.PFV-take $C_5$s-cloth = DEF **LOC** **$C_6$s-house = DEF** **inside**
*kį́* *ɔ́-dzɛ = ɛ*
give $C_1$s-woman = DEF
'He gave the cloth to the woman in the house.'/ 'He gave the cloth in the house to the woman.'

It is not considered grammatical to have two locative or temporal adverbials within a single core SVC. For instance in example (68), the first SVC in (a) with two temporal adverbials is considered ungrammatical. In the case of the SVC in (b) with two locational adverbials, the first adverbial phrase must refer to the location of the object rather than the location of the taking action.

(68) a. **a-kɔ̀* *kù̱-sà = a* ***kivòe***
$C_1$s.PFV-take $C_5$s-cloth = DEF **yesterday**
*kį́* *ɔ́-dzɛ = ɛ* ***òmonò***
give $C_1$s-woman = DEF **today**

b. *a-kɔ̀* *kù̱-sà = a* ***ní*** ***ke-pe = a*** ***mɛ̀***
$C_1$s.PFV-take $C_5$s-cloth = DEF **LOC** **$C_6$s-house = DEF** **inside**
*kį́* *ɔ́-dzɛ = ɛ* ***ní*** ***ɔ̀-nyɔ-nɔ̀*** ***mɛ̀***
give $C_1$s-woman = DEF **LOC** **$C_2$s-farm = DEF** **inside**
'He gave the cloth (which is) in the house to the woman on the farm.'

In an RRG analysis, the fact that aspect can be independently marked in each verb phrase shows they form separate nuclei. Since there can be only one locational

or temporal modifier, there is only one periphery. This combination suggests these SVCs are formed via core cosubordination.

### 3.4.3 Sequential SVCs in Avatime

SVCs combining sequential actions constitute the most semantically and morpho-syntactically divergent group in Avatime. They are the only SVCs which can be paraphrased with coordinated sentences, for instance (69) and (70).

(69) a. *mà-dɔ* *Gbàdzɛmɛ̀* *à-ba* *Ɔ̀vanɔ̀*
1s.PFV-move.from Gbadzeme SVM.1s.PFV-come Vane
'I left Gbadzeme came to Vane.'

b. mà-dɔ Gbàdzɛmɛ̀ **lɛ̌** **mà**-ba Ɔ̀vanɔ̀
1s.PFV-move.from Gbadzeme **and** **1s.PFV**-come Vane
'I left Gbadzeme and came to Vane.'

(70) a. *ma-tsà* *tomatoes=ye* *a-kpɛ* *ní* *kè-zi=a* *mɛ̀*
1s.PFV-cut tomatoes=DEF SVM.1s.PFV-put LOC $C_6$s-bowl=DEF inside
'I cut tomatoes and put them in the bowl.'

b. *ma-tsà* *tomatoes=ye*
1s.PFV-cut tomatoes=DEF
***lɛ̌*** ***mà-kpɛ*** *ní* *kè-zi=a* *mɛ̀*
**and** **1s.PFV**-put LOC $C_6$s-bowl=DEF inside
'I cut the tomatoes and put them in the bowl.'

It is also possible to modify each verb phrase in a sequential SVC with different temporal or locational adverbials, as in (71). This is not possible with other types of SVCs. So the separate adverbials in (72) force a sequential rather than an instrumental interpretation.

(71) a. *mà-dɔ* *Gbàdzɛmɛ̀* ***kivòe*** *à-ba*
1s.PFV-move.from Gbadzeme **yesterday** SVM.1s.PFV-come
*Ɔ̀vanɔ̀* ***òmonò***
Vane **today**
'I left Gbadzeme yesterday and came to Vane today.'

b. *ma-tsà* *tomatoes=ye* ***ní*** ***lị-vlɛ=lɛ̀***
1s.PFV-cut tomatoes=DEF **LOC** **$C_3$s-morning=DEF**
*a-kpɛ* *ní* *kè-zi=a* *mɛ̀* ***áblà***
SVM.1s.PFV-put LOC $C_6$s-bowl=DEF inside **now**
'I cut the tomatoes in the morning and put them in the bowl now.'

(72) a. *a-kɔ̀* *kà-wɛ=a* ***ní*** ***ke-pe=a*** ***mɛ̀***
$C_1$s.PFV-take $C_6$s-axe=DEF **LOC** **$C_6$s-house=DEF** **inside**
*tsà* *ò-se=lò* ***ní*** ***lị-ŋwàfụ=nɛ*** ***mɛ̀***
cut $C_2$s-tree=DEF **LOC** **$C_3$s-forest=DEF** **inside**
'He took the axe from inside the house, and cut the tree in the forest.'
Not: 'He used the axe (taken from inside the house) to cut the tree in the forest'

*b.* *a-kɔ̀* *kà-wɛ=a* ***kivòe*** *tsà* *ò-se=lò* ***òmonò***
$C_1$s.PFV-take $C_6$s-axe=DEF **yesterday** cut $C_2$s-tree=DEF **today**
'He took the axe yesterday, cut the tree today.'
Not 'He used the axe yesterday, cut the tree today.'

In terms of RRG, the fact that each verb can be modified by a locational or temporal modifier suggests sequential SVCs are formed by core coordination rather than cosubordination. In core coordination, each verb has its own periphery which allows them to be individually modified by locational and temporal modifiers, yet both verbs still share clause level markers such as mood (Van Valin 2005).

This property also means Avatime sequential SVCs do not have the macro-event property. Bohnemeyer and colleagues (2007) introduced the macro-event property as a way of evaluating whether constructions refer to single macro-events or not. A construction has the macro-event property if temporal operators necessarily scope over the entire construction (Bohnemeyer et al. 2007:497). SVCs are typically claimed to refer to single macro-events (e.g. Aikhenvald 2006; Comrie 1995; Durie 1997). Indeed in many West African languages there are other morpho-syntactic differences between constructions which allow independent modification by temporal and locational adverbials and SVCs which do not (Ameka 2003). This is not the case in Avatime. Bisang (2009) suggested just such a possibility when he suggested some action sequences may allow independent modification by temporal

adverbials but still refer to culturally determined event units and be described using SVCs. This appears to be the case in Avatime, where SVCs can only join action sequences which are understood by Avatime speakers to form coherent units (Section 3.3).

### 3.4.4 Avatime SVC subtypes discussion

The distinction between the sequential and other SVCs in Avatime is an example of a type of division commonly made within serialising languages. This type of division is known by many different names, such as symmetrical versus asymmetrical (Aikhenvald 2006), with different terms often being used for different languages or language groups. For instance, linking versus modifying has been used with West African languages (Bamgbose 1974), chaining versus integrated more specifically with Akan (Hellan, Beermann & Andenes 2003; Osam 1994), and narrative versus compact has mainly been used with Austronesian and Papuan languages (Pawley 2008; van Staden & Reesink 2008). The different terms have also been defined in different ways. Symmetrical SVCs are defined as those where all verbs come from open classes and have equal status in the construction (Aikhenvald 2006:22). Linking and chaining SVCs are defined as those which can be derived from (Bamgbose 1974:18) or paraphrased by (Osam 1994:195) coordinated clauses. Narrative SVCs are defined as those expressing a sequence of loosely integrated events (Pawley 2008:174) where the verbs can be independently modified by locational and temporal modifiers (Pawley 2008:174; van Staden & Reesink 2008:30). All these properties co-occur in Avatime sequential SVCs. Indeed, there appears to be a common idea behind all of these distinctions, separating the more coordinate-like SVCs used for combining sequential actions from the more modifying type SVCs and the clustering of these properties has been noted before in other languages (e.g. Pawley 2008).

The nuclear versus core SVC distinction is commonly discussed in descriptions of Austronesian and Papuan languages (e.g. van Staden & Reesink 2008) but not West African languages. When Foley and Olson (1985) introduced the distinction, they claimed it would not be relevant for most West African languages excepting the verb final languages Ịjọ and Igbo. More recent work on Austronesian languages has shown nuclear SVCs do in fact occur in verb medial languages with objects occurring between the verbs (e.g. Bril 2004; Crowley 2002; Durie 1997). Despite this discovery, the relevance of the distinction has not been reconsidered for West

African languages. The distinction between nuclear and core subtypes in Avatime is subtle. While it is not typical for the two types to be distinguished so subtly, it has been reported. For instance, in the Austronesian language Taba, the main implication of the nuclear versus core SVC distinction is how it affects verbal animacy restrictions (Bowden 2008). It is possible the nuclear versus core distinction will turn out to be relevant for more West African languages, and may even help explain some of the differences noted among SVCs within individual languages, for instance the differences in subject marking in Akan SVCs (Ameka 2003; Osam 1994). Indeed, a critical reading of Kießling's (2011) description of SVCs in Isu (West-Ring, Grassfields, Niger-Congo) suggests they may also be divided into nuclear and core subtypes.

The differences between the Avatime SVC subtypes are only observable in some conditions, such as when certain modifiers are used. This is not an unusual situation and has also been noted in other languages, for instance the situation in Avatime is quite similar to that of the SVC subtypes in Lao (Enfield 2007). It does, however, mean the subtypes cannot be used functionally by Avatime speakers to modify the meaning of a construction. Instead, the pairing of functions with subtypes is typically fixed with particular functions only being expressed using SVCs of a single type. This is especially true for the sequential SVCs. The nuclear and core subtypes have more functional overlap. Each subtype has its principle members (Table 2): modifying SVCs all appear in the nuclear subtype and argument adding SVCs are core SVCs. Other functions, such as 'take' SVCs used to mark themes and manner plus action SVCs, can occur in either type. These kinds of SVCs do not behave like nuclear SVCs in some ways and core SVCs in others. Rather, speakers appear to select one type or the other and then respond to all questions in a way consistent with that type. The two subtypes are thus distinct with some functions appearing in both types rather than forming a middle point on a cline. These cases where certain functions can be performed by SVCs of different types suggest the distinctions between the subtypes are not solely due to the semantics.

## 3.5 Conclusion

Avatime lies in the heart of the Kwa serialising area of West Africa, but it is also one of the Ghana-Togo Mountain languages known for their typological divergence from the surrounding Kwa languages. The description presented here shows Avatime

SVCs conform to the West African type (Ameka, 2003) in many ways. For instance, the subject must be a shared argument of all verbs and individual verbs can be marked for focus within the SVC. However, there are also ways in which Avatime SVCs are more like those from further afield. For instance, the reduced subject agreement markers have closer parallels in serialising languages outside rather than inside the West African region. The characteristics of the subtypes also have much in common with subtypes described among Austronesian, Papuan, and South-East Asian serialising languages and the literature on those languages helps inform an analysis of Avatime. This description of Avatime SVCs thus contributes to a better understanding of the range of variation within West African SVCs and also the possible similarities and connections between SVCs in different linguistic areas.

In this chapter, I have shown that SVCs form a distinct construction type in Avatime. I have described their morphosyntactic properties and semantic functions, and shown how they divide into three subtypes. The rest of this thesis builds upon this description using three different approaches to investigate the relationship between these SVCs and conceptual event units.

# 4 Do serial verb constructions describe single events? A study of co-speech gestures in Avatime

A version of this chapter appears as:
Defina, Rebecca. In press. Do serial verb constructions describe single events? A study of co-speech gestures in Avatime. *Language*.

## 4.1 Introduction

Serial verb constructions (SVCs) are often said to refer to single conceptual events. In fact, this property is commonly listed as one of their defining features (e.g. Aikhenvald 2006; Comrie 1995), and sometimes singled out as the primary feature from which others are derived (e.g. Bisang 2009; Durie 1997). However, this connection between SVCs and single events is not without its problems. Some propose an opposing view, whereby SVCs refer to multiple events, in contrast to single lexical verbs (Pawley 1987; Pawley 2011) or coverbs (Baker & Harvey 2010). However, a more fundamental issue has been the lack of clarity regarding what it means for a construction to refer to a single event and how one could test it (e.g. Crowley 2002; Foley 2010a; Pawley 2011; Senft 2008).

This chapter contributes towards a solution to this issue with the introduction and trial of a new method using the alignment of co-speech gestures to investigate conceptual event structure. Co-speech gestures are movements people make while speaking. They are produced frequently and often unconsciously. Since gestures visually express aspects of the conceptual message, they provide a window onto conceptual representations (e.g. Casasanto 2013; McNeill 1992). In addition, while they are temporally and semantically tightly connected to speech, they are produced independently from it (e.g. de Ruiter 2000; Kita & Özyürek 2003). This makes them an excellent tool for investigating the event structures referred to by different syntactic constructions. Specifically, constructions referring to multiple events should occur with multiple separate event gestures, while those relating to single events should only occur with single event gestures.

The alignment of gestures with SVCs and other multi-verbal constructions was tested in the serialising language Avatime. While other multi-verbal constructions frequently occurred with multiple event gestures, SVCs never did. This suggests that Avatime SVCs are indeed used to refer to single events. Such a finding is relevant not only for SVCs, but has broader implications for understanding and investigating relationships between conceptual and linguistic event structure. In the remainder of

this introduction I will briefly review the previous attempts to evaluate the event structure of SVCs, and introduce co-speech gestures, before discussing the specific hypotheses for the present study.

### 4.1.1 SVCs and events

As mentioned in Section 1.5, many researchers have claimed a connection between SVCs and single events (e.g. Aikhenvald 2006; Bisang 2009; Comrie 1995; Durie 1997). However, attempts to investigate this connection have suffered from the notorious difficulties concerning the relations between linguistic constructions and event representations, with many expressing dissatisfaction with the methods applied (e.g. Crowley 2002; Foley 2010a; Senft 2008). Previous evaluations of whether SVCs refer to single events have relied on five methods: intuition, translation, intonation, cultural restrictions, and the scope of temporal modifiers. I discuss these in turn below.

The most common approach is to use a consultant's or the linguist's own intuitions on what constitutes a single event (Crowley 2002; van Staden & Reesink 2008). Relying on the linguist's own intuitions can be problematic as they may not be sufficiently familiar with the language and culture in question. More generally, relying on intuitions is essentially problematic since deciding whether or not something constitutes a single – possibly complex – event or a collection of separate events is notoriously difficult and people often have different intuitions based on the same facts (e.g. Casati & Varzi 2008; Foley 2010a; Pawley & Lane 1998; Schultze-Berndt 2000:36–37; Schwartz 2008).

Another commonly used approach relies on translation (van Staden & Reesink 2008). Here a construction is said to refer to a single event if it can be translated into another language – typically English – using a single clause with a single finite verb. A more nuanced version, as employed by Baker and Harvey (2010), compares semantic structures of SVCs with those of single lexical verbs across many languages. In both cases, the core assumption is that single events in one language can be equated with single lexical verbs in another language. This is problematic for several reasons. Firstly, we know languages differ in the concepts they lexicalise in verbs (e.g. Jackendoff 1990; Malt & Majid 2013; Talmy 1985). So the results may differ depending on which language(s) are chosen for reference. Even within individual languages, the equation of single events with verbs is problematic. For

while some have suggested such an equation (e.g. Barsalou 1992; Croft 1990; Folli & Harley 2006; Parsons 1990), it has not been shown. There are also many who argue single lexical verbs can refer to multiple events, for instance the little *v* analyses of causative verbs like *feed* which are said to refer to distinct events of causing and eating (e.g. Hale & Keyser 1993; Pustejovsky 1991). At the very least, assuming an equation between verbs and events makes the discussion of whether or not particular linguistic constructions refer to single conceptual events circular (Givón 1991; Pawley 2011).

The third method, first employed by Givón (1991), uses pauses as evidence for the event structure of SVCs. Givón took Goldman-Eisler's (1968) finding that speakers pause when they are encoding the next unit of speech and concluded pauses in speech could be used as an indication of conceptual event boundaries. He examined three serializing languages of Papua New Guinea: Kalam, Tairora and Tok Pisin and found that the rate of pauses within SVCs was not distinguishable from the rate of pauses within single words. He concluded from this that SVCs must refer to single events. While this method is a great improvement over previous methods, it is also not without its problems. Firstly, it requires pain-staking phonetic analysis, so very few researchers have utilized it properly (Crowley 2002). There have also been doubts as to whether intonation units can demonstrate a non-circular relationship between syntax and event structure (Pawley 2011; Himmelmann 2013). Indeed, Himmelmann's (2013) recent study of intonation in a collection of texts across a range of serializing and non-serializing languages found that single intonational units at times included multiple clauses and events. He concluded from this that intonation units relate more to information structure than to either clause or event structure.

The fourth method for evaluating whether or not a construction refers to a single conceptual event focuses on cultural restrictions regarding what events it can describe (e.g. Bruce 1988; Diller 2006; Durie 1997; Enfield 2002). The productivity of SVCs is limited by the cultural notions of what constitutes a typical event. For instance, Bruce (1988:29) found it was possible to use an SVC to describe the commonly combined actions of climbing a tree and searching for insects in Alambak (Papuan), but not to describe the combination of climbing a tree and seeing the stars, which informants pointed out could be seen perfectly well from the ground. Researchers using this approach argue that since SVCs are limited by local cultural notions of legitimate event types, they must be restricted to refer to single –

culturally legitimate – event units. This method has the virtue of not relying on an assumed connection between linguistic structure and conceptual event structure. However, while it does show clear restrictions on the compatibility of the actions described by SVCs, it does not show these combined actions are necessarily conceptualized as single events.

Finally, recent work by Bohnemeyer and colleagues (2007; 2011) has introduced a new test of whether or not a construction refers to a single event: the macro event property. Constructions are said to have the macro event property if temporal modifiers scope over the whole construction. In contrast, constructions which allow multiple temporal modifiers with independent scope are said to refer to different macro events. This neatly captures the idea that single events should form a single coherent unit of space-time (e.g. Quine 1985; Zacks & Tversky 2001) However, it is a syntactic property corresponding to the unit of the core in Role and Reference Grammar (Bohnemeyer & Van Valin 2009) and it is not yet clear to what extent it reflects cognitive event representations (Bohnemeyer et al. 2007).

The results of these five methods do not always align. For instance, Baker and Harvey (2010) and Givón (1991) both set out to test whether SVCs referred to single events in two different cross-linguistic samples. Baker and Harvey used the translation method and concluded that SVCs refer to multiple events, while Givón used intonation and concluded that they refer to single events. Even when applied to the one language by the same researcher, the five methods can yield different results. For instance, in Avatime, intuition, intonation, and cultural restrictions point towards all SVCs referring to single events: There are cultural restrictions on the actions which can be combined using SVCs (Chapter 3) and they appear to occur within single intonational units and refer to single events. However, SVCs describing sequential actions cannot be translated into English using a single verb. These SVCs, all describing sequential actions, also do not have the macro event property since they can be modified by independent temporal modifiers as in example (1) (See also Section 3.4.3).

(1) *kivòe* *mà-dɔ* *Gbàdzɛmɛ̀* *à-ba* *Ɔ̀vanɔ̀* *òmonò*
yesterday 1s.PFV-move.from Gbadzeme SVM.1s.PFV-come Vane today
'Yesterday I left Gbadzeme, today I came to Vane.' (Elicitation_100712_AB)

While there are specific issues with each method, a central problem lies in the vagueness of the term *single event* and what it means to refer to one. The issue of event individuation has been much discussed among linguists, philosophers, and psychologists alike (e.g. Casati & Varzi 1996; Davidson 1969; Parsons 1990; Shipley & Zacks 2008). The question most relevant for SVCs is when a collection of actions constitutes a single event unit versus a collection of multiple events. Recent work by Zacks and colleagues helps shed light on this issue. In one study, they compared segmentations of events at different levels of granularity confirming that the boundaries of finer level events aligned with those of coarser level events (Zacks, Tversky & Iyer 2001). This alignment supports the idea that events are segmented hierarchically with each event containing sub-events and itself being a part of a larger event. For instance, *crossing the road* is made up of the sub-events of *checking for oncoming cars*, *stepping down off the curb* etc. and is itself a sub-event in the larger event unit of *walking to work*. Moreover, subsequent neurophysiological studies suggest people attend to multiple levels in this hierarchy simultaneously (e.g. Sharp & Donaldson 2007; Zacks et al. 2001; Zacks, Swallow, et al. 2006). This makes the distinction between single and multiple events essentially problematic. For while there are clearly some collections of actions which do not constitute single events – such as Apollo 11 landing on the moon and me eating breakfast this morning – the vast majority of single events have a dual nature and may also be seen as a collection of sub-events. This distinction is then more a matter of perspective and which level of the hierarchy is in focus at a particular time. The act of reference thus takes on an important role as the point of choosing which level of the event hierarchy to focus and the way of communicating that choice to your addressee. This is not a new idea and was a large part of Levelt's (1989) foundational account of the conceptual message and the processes underlying speech production. Thus, in order to investigate whether an SVC refers to a single event, we need to investigate the event structure of the conceptual message which this particular SVC is expressing to see whether it is framed as a single event or as a collection of events. This is what Givón (1991) aimed to achieve with his investigation of pause placement in SVCs. However, recent work on the nature of co-speech gestures suggests they may be a more informative and suitable tool.

### 4.1.2 Co-speech gestures: A window onto thinking-for-speaking

Co-speech gestures are meaningful visible movements of the hand, body, or face produced in connection with speech (Kendon 1986; Kendon 2004; McNeill 1992). They are tightly connected to speech and the two are often claimed to work together to form a single complex meaningful utterance (e.g. Enfield 2009; Kendon 2004; McNeill 1992). Among the types of gestures that have been distinguished, it is the ICONIC gestures which are of interest for the present study. Since these are the gestures which imagistically represent some part of the semantics of the utterance (McNeill 1992), and so may express aspects of the event.

Gestures typically consist of four phases: a PREPARATION, where the articulators are moved into position for the gesture; a STROKE, the main movement of the gesture; an optional HOLD phase, where the articulators are held still, generally immediately before or after the stroke; and a RETRACTION, where the articulators move back into a neutral or rest position (Kendon 2004; McNeill 1992). The strokes of iconic gestures are produced with a close temporal connection to their spoken affiliates (e.g. Butterworth & Beattie 1978; Kendon 1972; Krauss, Chen & Chawla 1996; McNeill 1985; McNeill 1992; Morrel-Samuels & Krauss 1992; Nobe 2000). This synchrony is maintained even when the speech is disfluent (Mayberry & Jaques 2000). One of the ways this is done is via the use of pre- or post-stroke holds which delay or extend the stroke of a gesture when the speech speeds ahead or lags behind gesture articulation (de Ruiter 2000; Kita 1990; McNeill 1992).

This temporal alignment with speech is one of the reasons gestures are frequently likened to intonation (e.g. Kendon 1980; McNeill 1992). Some have even suggested gesture and intonation are isomorphic (e.g. Hübler 2007). If so, using gestures to study event segmentation would be equivalent to using intonational units. However, recent research suggests that while gestures and intonation are tightly connected, the relationship is more complex. Several connections have been established. For instance, the peaks of gestural strokes tend to coincide with pitch accented syllables (Esteve-Gibert & Prieto 2013; Loehr 2012). Also whole gestural phrases – a stroke combined with its preparation, hold, and retraction – are loosely correlated with intermediate phrases (Loehr 2012). Intermediate phrases are a type of intonational unit smaller than the intonational phrase. They are bounded by pitch accents and inter-word junctures rather than the boundary tones and utterance final pauses which separate full intonational phrases. There are, however, also several

places of divergence. For instance, no correlations have been found at the level of gestural phases – such as the strokes as used in this study – or at the level of intonational phrases – as used in previous intonational studies of SVCs. The documented correlations in fact suggest that multiple gestural strokes could occur within single intonational phrases since these can have multiple pitch accents and intermediate phrases (Ladd 1996). Of course, not all of these strokes would necessarily relate to the event. Some may relate to an object or have a more deictic or pragmatic function and it is here that the extra semantic information available in gesture becomes incredibly valuable. In sum, both intonation and gesture relate to speech and so tend to align to a large extent. However, they do not align perfectly suggesting they relate to speech independently.

In addition, while gestures are indeed produced in tight connection to speech, mounting evidence suggests they are produced separately. For instance, several studies have found they can encode semantics not included in the spoken utterance (Kita & Özyürek 2003; Goldin-Meadow 2003). In fact, times when children produce gestures and speech representing different information have been identified as prime indicators of readiness to learn (e.g. Goldin-Meadow 2003; Pine, Lufkin & Messer 2004). For example, Pine et al.'s (2004) study examined children's explanations during a balancing task. Approximately one-third of the children produced mismatching gestures and speech at the beginning of the session, for instance producing gestures describing the relative weight of each side of the beam, but talking about the need to place the middle of the beam on the fulcrum. They found these children were the ones who benefitted most from instruction and showed the most improvement by the end of the session. Goldin-Meadow (2003) has suggested these mismatches indicate the children are beginning to entertain new concepts and explanations which they aren't yet able to verbalise, demonstrating a dual conceptualization stage in between shifting from one way of understanding to another. These mismatches also show that the encoding of conceptual representations into gesture and speech deviate at a very early stage of production. The likely relationship between speech and gesture production is diagrammed in Figure 4.1.

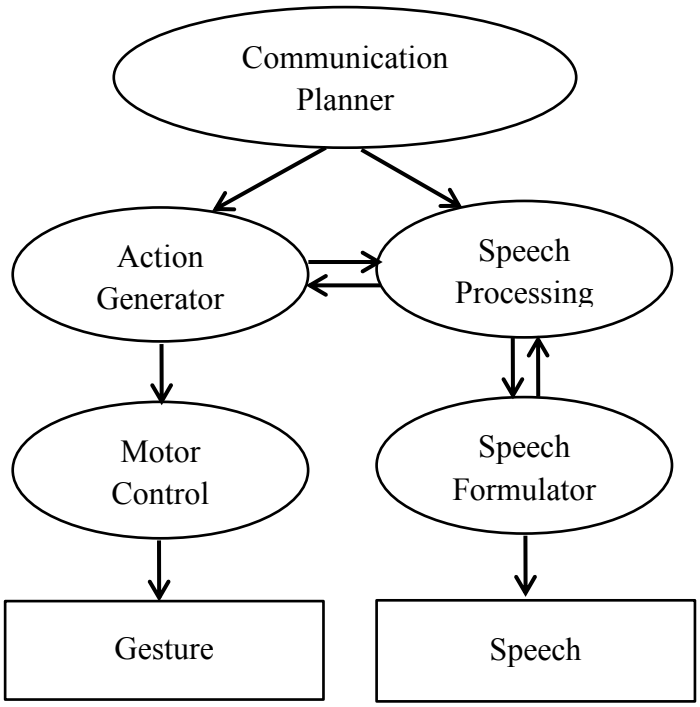



Figure 4.1: Diagram of likely relations between speech and gesture production processes, adapted from Kita & Özyürek (2003) and Chu & Kita (2009).

This relationship between gesture, speech, and conceptual message is what makes co-speech gestures such an excellent tool for investigating preverbal conceptual messages[1] and THINKING-FOR-SPEAKING (Slobin 1987; Slobin 1996a), and more studies are utilizing them for this purpose (e.g. Casasanto & Jasmin 2012; Goldin-Meadow 2006; Goldin-Meadow & Beilock 2010; Gullberg 2011; McNeill & Duncan 2000; Parril, Bergen & Lichtenstein 2013). Of particular note for this study, several researchers have claimed the level of cohesion during conceptualization is shown by whether a complex idea is expressed with a single gesture or broken up into multiple gestures (e.g. Kita & Özyürek 2003; McNeill & Duncan 2000). For

[1] Note this does not assume the preverbal conceptual message is a fully formed coherent unit before gesture and speech production begins. Indeed, the conceptual message appears to be formed incrementally in parallel with speech and gesture production (see Norcliffe & Konopka 2015). Whatever information is available at the conceptual message level would then be sent to both gesture and speech production systems in an incremental fashion.

example, Kita and Özyürek (2003) studied the gestures occurring with descriptions of a motion event by speakers of English, Japanese, and Turkish. While English speakers typically expressed rolling manner and downwards path in a single clause using a verb plus satellite combination like *roll down*, Japanese and Turkish speakers expressed them with different verbs in separate clauses, as in *descended the hill while rolling*. Speakers of all three languages frequently produced gestures conflating manner and path reflecting their simultaneity in the stimulus event. However, while these were often the only gestures English speakers produced, Japanese and Turkish speakers tended to combine them with additional separate gestures for manner or path. Kita and Özyürek argued that Japanese and Turkish speakers were led to encode manner and path separately by the way they packaged the information syntactically across multiple clauses. In a latter study, they found English speakers could be induced to produce separate manner and path gestures by asking them to describe the event using separate manner and path verbs in multiple clauses (Kita et al. 2007). This shows that the gestural patterns were indeed reflective of particular linguistic encoding choices rather than more long-term habitual encoding strategies.

### 4.1.3 Hypothesis

This study uses the alignment of co-speech gestures to test the hypothesis that Avatime SVCs refer to single events, in contrast to other types of complex clauses. Focusing on the alignment rather than the semantics of the gestures has two benefits. It avoids problems caused by the tendency for gestures to express only some aspects of situations. Also the measurement of temporal alignment is simpler and less subjective than that of gestural semantics, which means this method is more likely to be generally usable. Since the strokes of iconic gestures are produced in synchrony with the speech they relate to, the way iconic gestures relating to events – hereafter referred to simply as event gestures – overlap with speech can be used as a way of investigating conceptual event structures of particular syntactic construction types. There are two relevant event gesture-speech alignment patterns: Single Event Gesture in Total Overlap and Multiple Event Gestures. If a single event gesture is produced in overlap with all verbs in a construction, it suggests the verbs refer to a single event in the conceptual message. In contrast, if there are separate event gestures overlapping with different verbs within a construction, it suggests the verbs refer to different events in the conceptual message and so the whole construction is viewed as referring to a combination of multiple events. It is, of

course, also possible that there will be no event gestures overlapping a construction, or one event gesture which does not overlap all verbs. Neither of these situations provides definitive information about the event structure of the construction, as not all events are gestured and sometimes a gesture is not completely aligned with the spoken correlate. So if there is only one event gesture and it does not overlap all verbs it could have a narrow reference to one subevent or it could refer to an overarching event and be slightly out of alignment. In some cases the gestural semantics can help identify the latter of these two possibilities. The specific predictions for this study are thus that: (i) single event gestures overlapping multiple verbs will occur with Avatime SVCs, but not with other complex clauses, and (ii) separate distinct event gestures overlapping individual verbs will occur with other complex clauses, but not with SVCs.

## 4.2 Method

### 4.2.1 Data

Four native speakers of Avatime – three women and one man – aged 65-85 were invited to tell a story or describe how something was done. The resulting narratives and procedural descriptions totalled 42 minutes and 49 seconds and contained 646 utterances, where utterances were taken to be semantically and syntactically cohesive stretches of speech between pauses. To encourage naturalistic performances there was always an audience of native Avatime listeners present. These listeners were often familiar with some aspects of what they were being told, but never with the details. There was no mention of gesture before or during the recordings. All monologues were video and audio recorded. Table 4.1 summarizes the details of the monologues used.

### 4.2.1 Coding

All of the monologues were transcribed in Avatime and translated into English with the aid of native Avatime speakers also fluent in English. They were then glossed[2], before further syntactic and gestural coding. All annotation and coding was done using ELAN (Wittenberg et al. 2009).

[2] The Folkstory text was glossed by Saskia van Putten.

Table 4.1: The type, topic, length, and number of utterances contained in the four recordings used for this study.

| TYPE | TOPIC | LENGTH | TOTAL UTTERANCES |
|---|---|---|---|
| Procedural | Rice farming | 7min 5sec | 92 |
| Procedural | Midwifery | 7min 46sec | 192 |
| Narrative | Folkstory: Why we sacrifice rams at funerals | 7min 38sec | 149 |
| Narrative | Avatime migration history | 20min 20sec | 213 |

Instances of SVCs and other multi-verbal constructions were noted and coded as either SVC or OTHER COMPLEX. The distinctions between other complex constructions were collapsed in the main analysis, but are discussed later in Section 4.3.2. Utterances consisting of only one clause with a single verb were coded as SIMPLE, and those without a finite verb were coded as NONFINITE.

All iconic gestures relating to the event were tagged. These are gestures which imagistically represent some part of the described event. Gestures relating only to objects or locations were not included. For instance, a gesture produced with both hands flat and palm down starting in front of the chest and moving to the sides would be included if it occurred in a description of a flattening or smoothing type event. The same gesture would not be included if the speaker was describing a flat object. It was also noted to which verbs, if any, the gesture was semantically related.

The gestural phases – preparation, stroke, hold, and retraction – were identified by examining the video frame-by-frame. Strokes were identified following McNeil (1992) and Kendon (2004) as the meaningful part of the gesture, where the most effort is, and where the articulator is tensed. Following Kendon (2004), complex strokes were allowed, for instance a set of quick repeated movements to convey shaking an object was annotated as a single stroke.

The type of overlap was then determined by what speech overlapped with the stroke plus any poststroke hold of a gesture. If it overlapped all verb roots in the construction, it was coded as SINGLE TOTAL OVERLAP. If there was one gesture and it overlapped with at least one, but not all, verb roots, it was coded as SINGLE PARTIAL OVERLAP. If no verb roots overlapped the gestural stroke plus hold, it was coded as NOT ON THE VERB. If there were multiple distinct strokes within a construction, the

overlap was coded as MULTIPLE. Finally, if there were no event gestures, the overlap was coded as NO GESTURE.

(2) *lɛ̌* *a-ya=lɛ* *e-dù=i*
XXXXXXXXXXXXXXXXXXXXX
then $C_1$S.PFV-separate=$C_3$S.OBJ SVM.$C_1$S.PFV-put.on.flat.surface=CM
'Then she shared it (the porridge) out' (literally: separated it put it down)
(Folkstory_110406_QM _01:14)

Figure 4.2: Screenshot of the gesture in example (2). The speaker moves both her hands forward and down while describing how the woman separates individual servings of porridge and puts them down.

A couple of examples will help to clarify the coding. For instance in example (2) the speaker uses an SVC[3]. She also produces a gesture relating to both verbs: moving both her hands forward and down to represent the act of separating out a serve of porridge and putting it down. The duration of the stroke is indicated with Xs. Since

[3] Videos of gesture examples discussed in the chapter can be found in the archive https://hdl.handle.net/1839/00-0000-0000-0021-E9FE-C@view, with the exception of examples 4, 9, and 10 as the speaker has not released this text for the public.

there is only one event gesture and it overlaps and with both verb roots, it is coded as SINGLE TOTAL OVERLAP.

A single utterance may combine more than one clause type. For instance, one or both clauses in a coordinated construction may itself be complex. In these cases, embedded SVCs and other complex constructions are also noted. If any event gestures occur during this complex utterance, they are first annotated with respect to the lowest level clause and then to the higher level units. For instance in example (3), an SVC *bedzì ba kèdea* 'they returned-came back' is coordinated with a simple clause *bɛna- batɔ bɛna ɔ̀kʊ́tɔɛ* 'they reached- some (of them) reached some place'. There were two event gestures produced. The first is a gesture with the left hand, flat and palm down, moving from the left side to the space in front of the speaker's body. It relates to the 'coming' motion and possibly also the 'returning' aspect, especially as previous gestures have represented motion away from the body. It overlaps all verbs in the SVC *bedzì ba kèdea* 'they returned-came back'. The second event gesture is also produced with the left hand, flat and palm down, the hand moves from in front of the speaker's chest down towards his lap relating to the 'reaching' event. It overlaps with the first attempt at the second part of the conjoined clause *bɛna* 'they reached' and the beginning of the rephrasing to add the explicit subject. The SVC has a single event gesture overlapping both verbs so it is coded as Single Totally Overlapping. The simple clause has a single event gesture overlapping its single verb, not counting the repeat in the self-repair. However, since simple single verb clauses were not the focus of this study, embedded simple clauses such as this one were not separately noted. Nevertheless, this gesture is still relevant for the overarching coordinate construction, which is coded as having Multiple Gestures with an event gesture on each component.

(3) *ńte mè sị̀ lɛ̌ be-dzì ba kè-de=a*
XXXXXXXXXXXXXXXXXX
LOC:like.that inside COMP and $C_1$p.PFV-return come $C_6$s-back=DEF

*lɛ̌ bɛ-na ba-tɔ bɛ-na ɔ̀-kʊ́-tɔ=ɛ*
XXXXXXXXXXXXXX
and $C_1$p.PFV-reach $C_1$p-INDEF $C_1$p.PFV-reach $C_2$s-place-INDEF=CM
'So they came back and some (of them) reached some place.'
(Avatime-history_110905_BB_11:10)

### 4.2.2 Reliability

In order to check the reliability of the coding, a randomly selected subsection of the data was independently coded by a second coder. In cases of disagreement the initial coding was maintained. Two minutes were randomly selected from each text (18.7% of the total data) and independent coders coded both the syntax and the gesture. The second coder for the syntax was Saskia van Putten, a linguist also working on Avatime, and the second coder for the gestures was Emanuela Campisi, a linguist with extensive gesture coding experience. For the syntax, a simple comparison between the two codings showed a high degree of convergence, Cohen's kappa = 0.90. Since the coding of gestural phases is notoriously variable (Nobe 2000), a direct comparison of the phase boundaries was not attempted. Rather, the type of overlap was compared so that two codings were treated as the same if they yielded the same type of overlap with the syntactic construction. The convergence rate for types of gestural overlap was Cohen's kappa = 0.61. According to Landis and Koch (1977), this is substantial convergence. Still it is worth considering the nature of the divergences in more detail. Consultation with the second gesture coder showed 87% of the differences were due to disagreements in whether a gesture was iconic as opposed to deictic or pragmatic. The distinction between these types of gestures is known to be difficult (Kendon 2004:103–104). For instance, it can often be hard to determine whether a gesture accompanying a motion event is better treated as a deictic indication of the goal of motion or as an iconic referring to the path of motion. The initial coding had more of a tendency to analyse borderline gestures as iconic rather than deictic or pragmatic. Including additional non-iconic gestures in the analysis should not increase the risk of falsely confirming the hypothesis. Since non-iconic gestures may not relate specifically to the 'event' element in the utterance, they may have different alignment patterns. Thus, including them could reduce the strength of the predicted alignment pattern but not falsely confirm it. Only 7% of the divergences in the coding were due to differences in the timing of the stroke or hold of a gesture.

## 4.3 Results

The total numbers of SVCs and other complex constructions and simple and nonfinite utterances are listed in Table 4.2, along with the numbers of each kind of gestural overlap for each construction type. Further discussion will focus in on the

Single Total Overlap and Multiple Gestures for SVCs and Other Complex constructions. It is, however, informative to first briefly consider the full picture. Here one can see that SVCs make up a substantial proportion of the utterances in the sample (18%), though the various Other Complex constructions are much more frequent as a whole. Complex constructions are much more frequent as a whole. Event gestures also appear to occur disproportionally frequently with SVCs – in 64% of cases – as compared to other construction types – 44% for Other Complex constructions and 33% for simple clauses.

Table 4.2: The number of each kind of gestural overlap occurring with each type of utterance. Single Total Overlap refers to cases where a single event gesture overlaps all verb roots. It is therefore not applicable to the non-finite utterances which do not include a verb. Single Partial Overlap refers to cases where a single event gesture overlaps some of the verb roots. It is then only applicable to the SVCs and other complex constructions which contain more than one verb. Multiple refers to cases where there are multiple event gestures overlapping the utterance. Not On Verb refers to cases where there is a single event gesture which doesn't overlap any verb roots. Finally, No Gesture refers to cases where there were no iconic gestures relating to the event during that utterance.

| | SVC | OTHER COMPLEX | SIMPLE | NONFINITE | *Totals* |
|---|---|---|---|---|---|
| SINGLE TOTAL OVERLAP | 60 | 13 | 84 | n/a | 157 |
| SINGLE PARTIAL OVERLAP | 31 | 93 | n/a | n/a | 124 |
| MULTIPLE | 0 | 55 | 1 | 1 | 57 |
| NOT ON VERB | 4 | 3 | 8 | 2 | 17 |
| NO GESTURE | 53 | 209 | 191 | 21 | 474 |
| *Totals* | 148 | 373 | 284 | 24 | 829 |

### 4.3.1 Comparison of SVCs with Other Complex constructions

As predicted according to the single event hypothesis, SVCs frequently occurred with Single Totally Overlapping Gestures but never with Multiple Gestures. In contrast, Other Complex constructions frequently occurred with Multiple Gestures but only rarely with Single Totally Overlapping Gestures. This difference in alignment patterns was statistically significant. Figure 4.3 shows the distributions

observed. Chi-squared tests with Yates' correction were used to test the relationships between gesture overlap type and construction type as well as with text type and individual speaker differences. There was no relationship found between types of gestural overlap and individual speakers $\chi^2$ (3, N = 128) = 1.63, $p$ = 0.65. Likewise, there was also no relationship between gestural alignment and the two text types (narrative versus procedural) $\chi^2$ (1,N = 128) = 0.90, $p$ = 0.34. There was, however, a significant relationship between gesture overlap and construction type $\chi^2$ (1, 128) = 85.09, $p < 0.001$. This shows the differences in gestural alignment patterns were related to differences in construction type, rather than differences between text types or individual speaker styles.

As discussed in Section 4.1.3, Single Partially Overlapping Gestures can generally tell us little about event conceptualizations as they can arise from multiple scenarios. However, it is worth noting that 6 of the 31 event gestures partially overlapping with SVCs were clearly semantically related to both verbs, whereas none of the 93 event gestures partially overlapping with Other Complex utterances were. This provides further evidence supporting the single event conceptualization for SVCs.

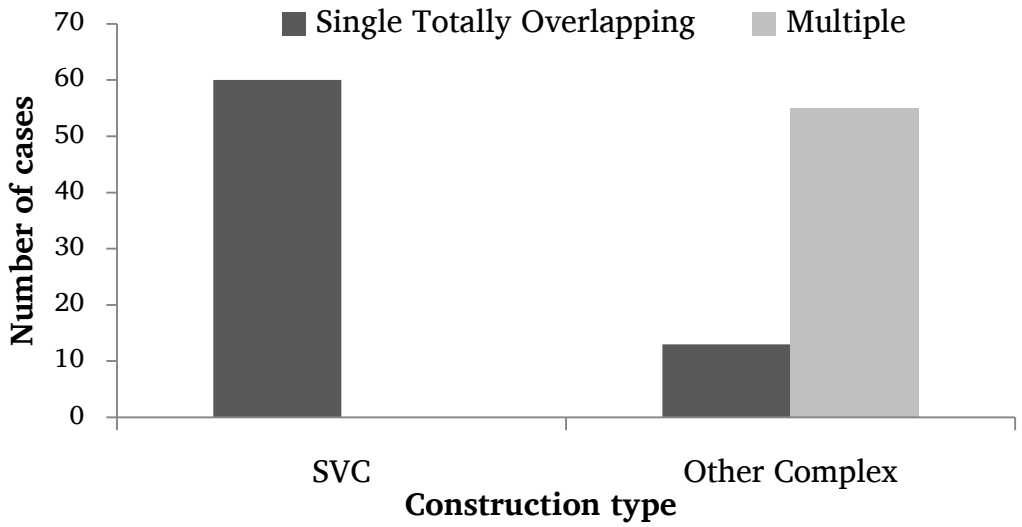



Figure 4.3: The number of Single Totally Overlapping and Multiple Gestures occurring with SVCs and Other Complex clauses. The difference between SVCs and Other Complex clause types is significant.

### 4.3.2 A closer look within each category

The hypothesis concerned the difference between SVCs and Other Complex clause types in general. However, neither SVCs nor the Other Complex clauses are

homogenous categories. Thus, in this section, I examine the various subtypes to determine whether the gestural alignment patterns were consistent within each category. There were too few cases to perform a statistical comparison among the subtypes, so I will discuss them qualitatively. I will begin with the SVCs, and then move on to the Other Complex clauses.

Within the SVCs, there is one subtype which stands out for potentially referring to multiple events. These are the SVCs referring to sequential actions, since they can be paraphrased with coordinate clauses in Avatime, are translated using two clauses in English, and do not have the macro event property (Section 3.4.3). There was only seven of this type of SVC in the present dataset. Three of these occur with a Single Totally Overlapping Gesture, as in example (4), and none occur with Multiple Gestures. This suggests that these sequential action SVCs tend to pattern like other Avatime SVCs and also refer to single conceptual events. Thus, even though these constructions allow multiple temporal modifiers and are translated into English with multiple clauses, Avatime speakers appear to focus on the larger event unifying these subevents, in this case of example (4) the event of gathering the rice to a central location.

(4) *kɔ bɛ-halị̀* *eh halị̀* *a-mụ=nà* *ɛ-manɔ̀* *ní*

XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

so $C_1$p.PFV-collect eh collect $C_3$p-rice=DEF SVM.$C_1$p.PFV-bring LOC

*eh ke-p… eh ní* *ke-tsripe=à* *mɛ̀* *ní* *ɔ̀-nyɔ=nɔ̀* *mɛ̀*

eh $C_6$s-h… eh LOC $C_6$s-clearing=DEF inside LOC $C_2$s-farm=DEF inside

‘So they collect the rice bring it to eh ho- to the clearing in the farm.’

(Rice-farming_100613_EN_05:26)

The Other Complex constructions consist of coordinated clauses, and subordinate clauses joined as adjunct, complements, or relative clauses. They also include nonfinite complement constructions where the matrix verb takes a nonfinite verb phrase as a complement. The numbers of Single Totally Overlapping and Multiple Gestures for each of these construction types are shown in Table 4.3.

Table 4.3: The number of Single Totally Overlapping and Multiple Gestures occurring with each type of Other Complex construction.

| | SINGLE TOTALLY OVERLAPPING GESTURE | MULTIPLE GESTURES |
|---|---|---|
| ADJUNCT | 0 | 28 |
| COMPLEMENT | 1 | 4 |
| RELATIVE | 3 | 7 |
| COORDINATE | 2 | 15 |
| NONFINITE COMPLEMENT | 7 | 1 |

Firstly, the adjunct constructions conform exactly to the pattern predicted for the Other Complex constructions: they occur with Multiple Gestures, but not Single Totally Overlapping ones. For instance, the adjunct construction in example (5) has three event gestures. The first is on the verb *dzi* 'return' which is cut short before the utterance is rephrased. The second gesture overlaps the whole rephrased matrix clause. The speaker spreads both hands out and around from in front of his torso to his sides, in reference to the spreading out and dividing of the Avatime people into the various towns. The final gesture occurs in the temporal adjunct clause introduced by the clause linkage marker *xé* and overlaps with the verb *kị* 'give'. It is a short thrust of the hands away from the speaker in reference to the giving action.

(5) *kɔɛ kui-dzi–*
XX
so 1p.PFV-return

*bɛ-lɛ̀ e-ku te petee pɔ́-ɛ*
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
$C_1$p.PFV-share SVM.$C_1$p.PFV-enter like.that all finish-CM

*xé àblé ké bɛ-kị ị̀-mà=nɛ è-nyi=nà*
XXXXXXXXX
CLM now same $C_1$p.PFV-give $C_2$p-town=DEF $C_3$p-name=DEF
'So we return– they divided throughout, before they gave the towns names.'
(Avatime-History_110905_BB _11:45)

The complement, relative and coordinate constructions all trend towards this pattern as well. They often occur with Multiple Gestures, where one event gesture overlaps the verb in the matrix clause and another overlaps the verb in the subordinate or coordinate clause. An example of this can be seen in the coordinated clauses shown in example (3) in Section 4.2.1. However, these construction types also sometimes occurred with single event gestures overlapping both verbs. These exceptional cases are discussed below.

The data included only one complement construction with a Single Totally Overlapping Gesture, shown in example (6). Here the speaker is describing a baby's birth and produces a single event gesture overlapping all verbs in the complement construction 'God helped me to deliver her like this'. Her hands form a V-shape with the fingers of both hands touching at the bottom. She moves this V-shape slowly from in front of her stomach down and towards her left, coming to rest on her left thigh. This gesture appears to relate to the delivery of the baby rather than the help from God, yet it is timed over the whole complement construction. This gesture is on the border between iconic and pragmatic. Thus one possible explanation for the unusual timing is that it is in fact a pragmatic gesture and does not refer specifically to the birthing event after all.

(6) *lì-po = lè* *gì* *ma-pɔ̀* *Pearl = i* *mawu = yɛ*

$C_3$S-time = DEF REL 1S.PFV-birth Pearl = CM God = DEF

*a-kɔ̀ = mɛ* *a-pɔnɔ̀* *sị̀* *me-bu = ye*

XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

$C_1$S.PFV-take = 1S.OBJ SVM.$C_1$S.PFV-help COMP 1S.PFV-remove = $C_{1S}$.OBJ

*ple* *te*

XXX

descend like.this

'When I brought forth Pearl, God helped me to deliver her like this.'

(Midwifery_110901_AB _04:40)

There were two coordinate clauses with Single Totally Overlapping Gestures, both produced by the same speaker. He is the oldest speaker in this data set, at 85, has a

rather slow gestural style, and also tends to produce a lot of coordinate constructions. It is possible these single overlapping event gestures are a part of his particular style. An example can be seen in (7). The speaker's right hand starts on his lap and moves up to trace a large arc to his right side ending near his head. It remains there, relaxed, until the next stroke a couple of clauses later. Here the gesture seems to focus on the higher order event of moving from one place to another, while the speech describes the subevents of getting up or leaving and then moving up to the new location. Alternatively, the gesture may represent the second subevent only even though it also overlaps with the rising verb.

(7) *lɛ̌ lị-kị bɛ-yɔ xé bɛ-trɛ e-mu ní*
XXXXXXXXXXXXXXXXXXXXXXXXXXXX
then $C_3$s.PFV-give $C_1$p.PFV-rise CLM $C_1$p.PFV-go SVM.$C_1$p.PFV-ascend LOC
'Then that forced them to rise and move up to-'
(Avatime-history_110905_BB_07:27)

Figure 4.4: Screenshot of the gesture in example (7). The speaker moves his right hand from down near his side up in an arc to above his head while describing how the Avatime people left the place they had been staying and moved up to a new place on higher ground.

Three relative clauses were produced with Single Totally Overlapping Gestures. Notably, they all provide further explication of the event, rather than the otherwise more common use of relative clauses to provide more information about one of the participants in the event. These cases, thus, have more of a single event reference, both in the gesture and in the semantics of the spoken utterance. In example (8), the speaker is describing how a baby will turn before it is born. She makes fists with both hands one on either side of her stomach and circles them around clockwise. This circling motion is repeated until the end of the full complex clause construction.

(8) *kɔ* *ɔ-nị̀vɔ = ɛ* *èé-tsyí* *te*
XXXXXXXXXXXXXXXXXXXXXXXXXXX
so $C_1$s-child = DEF $C_1$s.PROG-turn like.that

*kị́lɛ* *gì* *èé-tsyí* *te* *petee* *rrrr*
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
how REL $C_1$s.PROG-turn like.that all continuously
'So the baby will be turning how it is turning like that all along.'
(Midwifery_110901_AB_09:46)

The nonfinite complement constructions are the only Other Complex constructions which deviate from the general pattern. They behave more like the SVCs, with more Single Totally Overlapping than Multiple Gestures. For instance, the nonfinite complement construction in (9), where the speaker describes rice starting to drink water and produces a single event gesture overlapping the entire construction. She moves her right hand up from in front of her chest to in front of her shoulder and then dips it downwards towards herself as if she is taking in the water.

(9) *e-kpese* *kù-ni = o* *ɔ̀-ŋwɛ̀*
XXXXXXXXXXXXXXXXXXXXXXXX
$C_3$p.PFV-start $C_5$s-water = DEF INF-drink
'It (the rice) starts to drink water.' (Rice-farming_100613_EN_04:07)

The nonfinite complements are also syntactically and semantically very different from the rest of the Other Complex clauses (See Section 2.7.2.3.1). The subordinate

verb is in a nonfinite verb phrase, rather than a full finite clause. So there is only one full finite clause in these constructions. The matrix verb is semantically light and modifies the aspect or modality of the subordinate verb which provides the main event semantics. It is, therefore, not surprising that they tend to occur with Single Totally Overlapping Gestures and apparently describe single events.

## 4.4 Discussion

The results show a strong tendency for SVCs to occur with Single Totally Overlapping Gestures and for Other Complex clauses to occur with Multiple Gestures. This suggests SVCs do indeed refer to single conceptual events, while other complex clauses refer to multiple events.

Nonfinite complement constructions were the only Other Complex clauses which seemed to refer to single rather than multiple events. They are also unique among the Other Complex clauses as they take a nonfinite verb phrase rather than a full finite clause as a complement. The finding that Avatime SVCs and nonfinite complements tend to have Single Totally Overlapping Gestures supports the view that conceptual event units during thinking-for-speaking relate more to clauses (e.g. Evans 2010; Jackendoff 1991; Kita & Özyürek 2003; Pustejovsky 1991) than to individual verbs (e.g. Baker & Harvey 2010; Croft 1990; Folli & Harley 2006; Malaia 2014).

While there was a strong tendency for SVCs and nonfinite complements to occur with Single Totally Overlapping Gestures and Other Complex constructions to occur with Multiple Gestures, this tendency was not absolute. This suggests there is not a simple isomorphism between syntactic structure and conceptual event structure. Indeed, coordinated clauses should be the most likely to refer to multiple events, yet there were some cases where they occurred with Single Totally Overlapping Gestures, for instance example (7) above. Conversely, simple clauses with a single verb should be the most likely to refer to single events, yet there were cases where they occurred with multiple event gestures, such as example (10). Here the speaker is describing placing rice grains in the sun to dry. She first makes one gesture with both hands in front of her chest moving straight down towards her lap overlapping with the subject agreement and directional prefix on the placement verb. She then makes a second gesture with both hands starting from in front of her chest moving down and spreading out, overlapping with the object and location. The gestures

appear to describe placement and spreading out actions, motions required for placing rice in the sun though the second action is not explicit in the lexical description.

(10) *wɔ-zɛ-plɛ* *lá* *we=o-we=o* *kò*
XXXX XXXXXXXXXXXXXXX
2s-IT-put $C_3$s=LOC sun=DEF-sun=DEF just
'You put it (the rice) in the sunshine (moving in a direction away from the origin).' (Rice-farming_100613_EN_06:59)

This is in line with previous studies (e.g. Garber & Goldin-Meadow 2002; Kita & Özyürek 2003) which have highlighted cases where gesture and speech differ in the aspects of the conceptual message they encode. In the present case, it is an issue of one focusing on a higher order event unit while the other takes a more granular view focusing in lower order subevents.

## 4.5 Conclusion

Researchers working on SVCs have generally assumed they refer to single conceptual events (e.g. Aikhenvald 2006; Bisang 2009; Comrie 1995; Durie 1997). However, evidence to support this connection has been elusive. The study presented here is the first investigation of iconic event gestures occurring with SVCs. It shows SVCs in Avatime tend to occur with single gestures overlapping the entire construction. This suggests that they do, in fact, refer to single conceptual events. In contrast, other complex clauses in Avatime tend to occur with multiple distinct gestures, suggesting they refer to multiple events. This provides long awaited evidence supporting the assumed connection between SVCs and single events and offers a new method for evaluating the relationship.

# 5 Perceiving events: Influences of serial verb constructions and familiarity

## 5.1 Introduction

Whenever we perceive activity, we divide it into individual event units. This discretization of activity is done very quickly (Strickland & Keil 2011) and doesn't require deliberate attention (Kurby & Zacks 2008; Zacks et al. 2001). It also has a profound influence on the way we remember events (e.g. DuBrow & Davachi 2013; Newtson & Engquist 1976; Sargent et al. 2013; Swallow, Zacks & Abrams 2009). Previous research (e.g. Zacks, Tversky & Iyer 2001) has suggested the way we segment events may be influenced by our familiarity with them as well as the ways we describe them. These possible influences of familiarity and linguistic description are intertwined in the case of serial verb constructions (SVC), since they are often used to describe culturally recognisable or familiar events. On one hand a difference according to familiarity could motivate a difference in how events described with SVCs are perceived, on the other hand it could also mask an influence of the SVC itself. This chapter uses a new measure known as "dwell time" (Hard, Recchia & Tversky 2011) to investigate the possible influences of cultural familiarity and the use of serial verb constructions on event segmentation during perception.

One of the major findings from previous research on event segmentation is that events are segmented hierarchically (Zacks, Tversky & Iyer 2001), with each event consisting of subevents and itself contributing as a subevent to a larger unit. For example, the event of someone boiling water consists of subevents of putting water in a pot, turning the stove on, and waiting for the water to boil. It may also be a subevent in a larger event such as making tea or cooking dinner, see Figure 5.1.

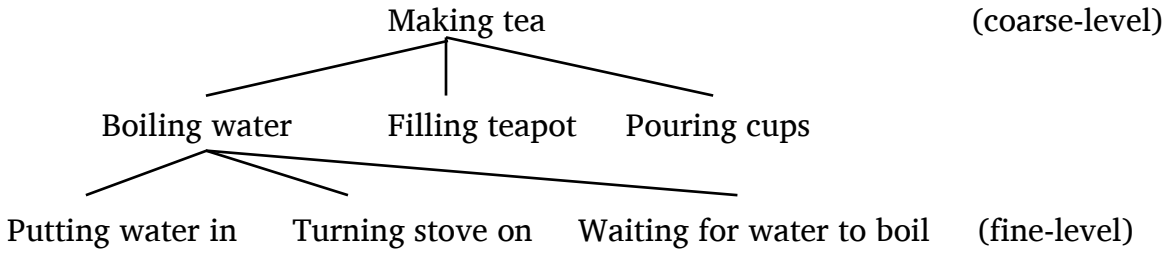



Figure 5.1: Schematic description of the hierarchical structure of an event, showing fine- and coarse-grained event units.

Previous studies of how we segment events during perception (e.g. Avrahami & Kareev 1994; Hard, Tversky & Lang 2006; Zacks 2004; Zacks, Tversky & Iyer 2001) have relied on the explicit segmentation method developed by Newtson (1973). According to this method, participants are asked to watch videos and report when they think one unit finishes and another begins. They are typically asked to segment each video twice: once at a fine-grain with smaller event units, and once at a coarse-grain with larger units.

These studies have identified several factors influencing the way we segment activity. Firstly, characteristics of the perceptual stimulus play a significant role. For instance points where the actor accelerates or changes direction are associated with boundaries between events, particularly at the fine-grained level (Hard, Recchia & Tversky 2011; Zacks 2004). Secondly, event boundaries frequently occur at points of goal-attainment (Baird & Baldwin 2001). Whether or not the perceiver understands the goals of the actor also matters: when people do not understand the goals they tend to segment activity more finely (Wilder 1978a; Zacks 2004). Thirdly, statistical knowledge about the tendency for particular actions to co-occur is utilized for segmentation (Avrahami & Kareev 1994). Finally, the way people segment events is influenced by task demands: if people are attending to characteristics of the actor rather than the actions, they tend to segment at a coarser level (Cohen & Ebbesen 1979); and when they are asked to explicitly describe actions, they segment in a more hierarchical manner with more agreement between fine and coarse-grained event boundaries (Zacks, Tversky & Iyer, 2001).

The influences of goal knowledge and task on event segmentation are in line with Zacks and colleagues' (2007) influential event segmentation theory. According to this theory, event segmentation is centred around the process of predicting future perceptual input on the basis of event schemata derived from previous experience. When witnessing activity, these event schemata are used to form a model of the event currently occurring. This event model drives predictions of future sensory input, which are then checked against incoming sensory information. Mismatches indicate the event model is no longer an accurate model for what is happening. This signals that the event model needs to be updated and this point of updating becomes a boundary between events.

This theory fits the reported effects. If the perceiver understands the actor's goals, they are more likely to form a well-fitting event model and thus encounter fewer

points of mismatch, resulting in a coarser segmentation (Zacks et al. 2007). Attending to the actor or the action also influences which event schemata are active and so what event models are constructed, in turn influencing the location of event boundaries (Cohen & Ebbesen 1979). Cohen and Ebbesen (1979) also suggested that collections of, rather than individual, event schemata are involved in event comprehension. This provides an explanation for the influence of description tasks on event segmentation: in order to describe an event, a particular event schema is chosen from amongst the several possibilities and this leads to a more consistent hierarchical organisation (Zacks et al. 2007).

The influence of event schemata on segmentation also suggests familiarity with an event would influence the way one segments it. If a person is familiar with an event, they should have a more detailed and accurate event schema leading to the formation of a more accurate event model. This would mean they would encounter fewer points of mismatch between the generated predictions and incoming information. They should, therefore, report fewer event boundaries and divide the activity into larger event units. In contrast, if someone is unfamiliar with the event, they will not be able to produce an accurate event model and so report more event boundaries and divide the activity into smaller event units. Such an influence of familiarity has been predicted previously (e.g. Hard, Tversky & Lang 2006; Heider 1958; Zacks et al. 2007; Zacks, Tversky & Iyer 2001). However, evidence to support it has been elusive.

An early study by Newtson (1973) showed that people begin to segment action more finely following an unpredicted interruption in an activity (the actor stopped assembling a model of a molecule and took off his right shoe and sock before rolling up his left pant leg and returning to the assembly). Follow-up work by Wilder (1978b) showed that people segment events more finely when activity begins unpredictably (actions in random order) versus predictably (actions in repeating logical order), and when the predictability of the actions changes, shifting either from predictable to unpredictable or vice versa. People will also segment activity more coarsely after viewing it several times than they do when viewing it for the first time (Hard, Tversky & Lang 2006). It has also been established that people segment events more coarsely when they understand the goals behind the actions (Wilder 1978a; Zacks 2004). While these issues of predictability, repeated exposure, and understanding of goals all relate to familiarity, they do not show whether general familiarity with an event leads to coarser segmentation.

So far, there has only been one study to directly compare the segmentation of familiar versus unfamiliar events (Zacks, Tversky & Iyer, 2001). In this study, Zacks and colleagues first collected ratings for frequency, familiarity, and knowledge of a set of 45 everyday activities. From these they selected two familiar activities (making a bed and doing the dishes) and two unfamiliar activities (assembling a saxophone and fertilizing houseplants) rated high or low on all three scales respectively. They then asked participants to segment videos of these four events into fine- and coarse-level units following the Newtson (1973) paradigm. They found no difference in the lengths of event segments for familiar versus unfamiliar events. There was greater alignment between fine- and coarse-level event boundaries among familiar events than unfamiliar events. However, this was not a robust effect and did not replicate in follow-up studies (Zacks, Tversky & Iyer 2001). Zacks and colleagues hypothesized that the unfamiliar events in their stimuli may not have been unfamiliar enough to yield a sufficient difference in familiarity and a clear difference in segmentation patterns. An alternative explanation is that human actions may be so strongly constrained by common physical laws and general patterns of goal attainment that familiarity is only able to exert a very small influence. The present set of experiments aims to test this further by comparing events from different cultures, thus providing a greater difference in familiarity.

The importance of event schemata for segmentation also suggests a relationship between language and perceived event units. Patterns in how we habitually describe events are likely related to the way we structure our event schemata and thus in turn to the way we segment them. For instance, habitual use of serial verb constructions (SVCs) to describe certain activities could be linked to event schemata which treat those activity sequences as single event units.

The possibility that SVCs could influence the way people segment events is intertwined with the issue of familiarity. It has often been noted that SVCs are only used to describe events which are culturally recognizable as coherent units and so familiar to the speaker and their interlocutors (e.g. Bruce 1988; Diller 2006; Durie 1997; Enfield 2002). While this restriction does not apply to many Avatime SVCs which serve a more grammatical function, such as adding an instrument or benefactive argument (Chapter 3, Section 3.3), it does apply when SVCs are used to describe sequential actions. Sequential actions can generally be described in Avatime using either coordinated clauses or an SVC. However, for an SVC to be

acceptable the actions must form a recognisable event type. For instance, the action sequences in examples (1) and (2) can be described using SVCs (1a & 2a) or coordinated clauses (1b & 2b). In contrast, the action sequence in (3) can only be described using coordinated clauses.

(1) a. *ɔ-dzɛ* *a-dɔ* *ɔ̀-ma=nɔ̀* *mɛ̀* *ba* *sku*
$C_1$s-woman $C_1$s.PFV-move.from $C_2$s-town=DEF inside come school
'The woman left the town, came to school.'
(The school referred to is a little outside of the town)

b. *ɔ-dzɛ* *a-dɔ* *ɔ̀-ma=nɔ̀* *mɛ̀*
$C_1$s-woman $C_1$s.PFV-move.from $C_2$s-town=DEF inside
*lɛ̌* *a-ba* *sku*
and $C_1$s.PFV-come school
'The woman left the town and came to school.'

(2) a. *ma-tsà* *tomatoes=ye* *a-kpɛ* *ní* *kè-zi=a* *mɛ̀*
1s.PFV-cut tomatoes=DEF SVM.1s.PFV-put LOC $C_6$s-bowl=DEF inside
'I cut tomatoes, put them in the bowl.'

b. *ma-tsà* *tomatoes=ye*
1s.PFV-cut tomatoes=DEF
*lɛ̌* *mà-kpɛ* *ní* *kè-zi=a* *mɛ̀*
and 1s.PFV-put LOC $C_6$s-bowl=DEF inside
'I cut tomatoes and put them in the bowl.' (Elicitations_2012)

(3) a. *ma-tsà* *tomatoes=ye* *lɛ̌* *mà-sɛ* *ke-pe=a* *mɛ̀*
1s.PFV-cut tomatoes=DEF and 1s.PFV-leave $C_6$s-house=DEF inside
'I cut tomatoes and left the house.'

b. **ma-tsà* *tomatoes=ye* *a-sɛ* *ke-pe=a* *mɛ̀*
1s.PFV-cut tomatoes=DEF SVM.1s.PFV-leave $C_6$s-house=DEF inside
Intended: 'I cut tomatoes, left the house.' (Elicitations_2012)

The link between familiarity and use of an SVC to describe a sequence of actions means that familiarity could present a confounding factor when investigating the possible interactions of SVC use and event segmentation. The experiments in this chapter thus set out to first investigate the potential influence of familiarity on

event segmentation and then the possible additional influence of SVC use. In order to maximise the variation in familiarity, a cross-cultural approach was adopted comparing event segmentations of Avatime and Dutch participants in Experiments 1 and 2. The prediction was that people would segment familiar events from their own culture more coarsely. Controlling for any effect of familiarity opens the possibility of investigating the role of SVCs in event segmentation more directly. Experiment 3 was a step in this direction. Avatime participants were primed with either SVCs or conjoined clauses before completing a segmentation task. It was hypothesized that priming with SVCs should lead to activation of event schemata with more unified event segments and thus coarser event segmentation. The data from Avatime participants for all three experiments were collected in a single field trip before the collection of the data from Dutch participants and before conducting any analyses.

## 5.2 General methods

As mentioned in the introduction to this chapter, previous studies of event segmentation have utilized the explicit boundary reporting method developed by Newtson (1973). This method of explicitly asking participants to report event boundaries has worked well with the typically English speaking university students participating in these studies. However, pilot testing of this method with Avatime participants was not successful. Several versions of the instructions were trialled in both Avatime and English, but participants did not feel confident that they understood what they were being asked to do and did not wish to continue with the task. There are several possible explanations for this. The task is likely easier for people who have followed certain styles of schooling or socialization where the breaking down and analysing of things into components is encouraged and practiced. Literacy skills may also be helpful. Difficulties in understanding the task may also be linked to differences in language, for instance Avatime has no good translation equivalent for ‘event’, so various alternatives such as ‘what the person is doing’ were used instead. Whatever the motivations underlying this difference are – and there are likely several working together – the result is that the traditional explicit event segmentation task is not viable for this broader sample of participants.

Since explicit segmentation was not an option, an alternative method had to be employed. A recent study by Hard, Recchia and Tversky (2011) provided such an

alternative. Hard and colleagues adapted the established self-paced reading paradigm (Just, Carpenter & Woolley 1982) commonly used to study language processing in order to study event segmentation. In self-paced reading studies, participants read a text one fragment at a time (e.g. word-by-word) and press a button in order to see the next fragment. The time between displaying the text and the participant's button press triggering the next fragment is recorded. The fundamental idea is that reading time increases at points with greater processing requirements. In particular, times are longer for fragments at the end of phrase boundaries where previous elements need to be integrated into larger units. This increase is modulated by the type of phrase boundary, so reaction times are longer at the ends of clauses and even longer at the end of sentences (Haberlandt & Graesser 1989). Hard and colleagues adapted this paradigm to the study of events by using slideshows of still images rather than text fragments. They made video recordings of events at a frame rate of 30 frames per second. They then extracted still images from these videos in the middle of each 1 second interval, so frame numbers 15, 45, 75, etc. Participants viewed the event slideshows one image at a time and pressed a button to trigger the next image. These reaction times were used to calculate what the authors referred to as the 'dwell times'. The prediction was participants should dwell longer at event boundaries since these points require greater processing and integration. Indeed, when these dwell times were compared with explicit segmentation judgements, they were found to peak at event boundary locations with the level of the peak varying according to the granularity of the boundary with longer dwell times occurring at coarser event boundaries (Hard, Recchia & Tversky 2011).

In another study, Meyer, Baldwin and Sage (2011) measured dwell times in young children, another participant group for which the explicit segmentation task is not viable. Meyer and colleagues compared dwell times of 3- and 4-year-old children with explicit event boundary judgements of experienced adult coders. Just as in the Hard et al. (2011) study, dwell times were longer at boundary points than at points within event units and coarse level boundary points had longer dwell times than fine level boundary points. The fact that this measure is implicit and works well even with young children suggests it is appropriate for measuring event segmentation across a broad range of participant groups.

Previous studies have compared dwell times to explicit event boundary judgements, either from the same participants or from a different participant group.

This was not possible with the current study, since the explicit segmentation task could not be done with the Avatime participants and Avatime segmentation patterns could not be assumed to correspond to Dutch segmentation patterns. Thus, this is the first study to attempt to use dwell times as the sole indicator of event boundaries. Since there was no established method for analysing this data, several different methods were employed to investigate what the dwell time data on its own could tell us about event segmentation. These are discussed in more detail in the Results section of Experiment 1 (Section 5.3.2)

## 5.3 Experiment 1

This experiment tests the hypothesis that people familiar with an event will segment it more coarsely than those for whom it is unfamiliar (e.g. Hard, Tversky & Lang 2006; Heider 1958; Zacks et al. 2007; Zacks, Tversky & Iyer 2001). Events were selected which were either very familiar among Avatime but not Dutch people, or, conversely, very familiar among Dutch but not Avatime people. Avatime and Dutch participants viewed both sets of events. The prediction is that Avatime participants will segment Avatime familiar events more coarsely than the Dutch participants, while Dutch participants will segment Dutch familiar events more coarsely than the Avatime.

### 5.3.1 Method

#### 5.3.1.1 Participants

Twenty-two Avatime participants, aged 12-16 (mean 13.5), were recruited at Amedzofe Junior High School, Ghana. Two participants were tested but excluded because they were clearly not attending to the task. Most participants had lived only within the Avatime traditional area. The only exceptions to this were two participants who were born in neighbouring regions and moved into the Avatime area at a young age. All Avatime participants were fluent in Avatime and also spoke Ewe and English, four additionally spoke Twi.

Twenty-one Dutch participants, aged 12-17 (mean 14.0), were recruited in the Netherlands at Pax Christi College Druten, Olympus College Arnhem, and via advertisements in Nijmegen. One Dutch participant was excluded for not attending to the task. Most participants had lived only in the Netherlands; only one was born outside the country and moved to the Netherlands at a young age. They were all

fluent in Dutch and also spoke English to varying degrees of fluency. Twelve additionally knew some French, eight some German, two some Latin, two some Spanish, one was fluent in Macedonian, and another was fluent in Portuguese.

#### 5.3.1.2 Materials

Four Avatime familiar events and four Dutch familiar events were selected, see Table 5.1[1].

In order to select the Avatime familiar events, 20 potential events were chosen based on my experience in the field. The criteria for these events were that they be performed by a single actor, take approximately one to five minutes, occur frequently in the Avatime community, and be unfamiliar to people in the Netherlands. These 20 events were filmed in a Likpe village with Likpe actors who would not be recognised by Avatime participants. The Likpe are another minority (Ghana-Togo Mountain language speaking) group located approximately 50km north of the Avatime area. Likpe villages have a similar appearance to Avatime villages and there is a great deal of cultural overlap. The 20 filmed events were then shown to three Avatime teenagers to judge their familiarity and whether there were any differences between how the Likpe and Avatime performed the event. Fourteen of the events were reported as highly familiar and performed in the local manner by all three judges. The four experimental items were selected from these fourteen highly familiar events based on the following factors: event structure – events with highly repetitive actions such as pounding fufu (a favoured starch based staple eaten with soups), weeding, and washing clothes were excluded; maximum differentiation from each other – events were chosen so that they were each enacted by a different actor, and only one of several fire-related events was selected; and length – very short events were excluded.

[1] Although eight is a relatively small number of items, it is more than commonly used in event segmentation studies which tend to feature four stimuli events or even fewer (cf. Hard, Recchia & Tversky 2011; Hard, Tversky & Lang 2006; Newtson 1973; Zacks & Tversky 2001; Zacks 2004). Each item provides a large amount of data. The measure is also likely to be compromised if participants become bored, so the task duration should be kept to a minimum.

Table 5.1: The experimental stimuli (four Avatime familiar and four Dutch familiar events), with an example still from each video, the duration of the video in minutes and seconds, and the number of slides used in the experiment.

| **Description of event** | **Example still image** | **Duration (mm:ss)** | **No. of slides** |
|---|---|---|---|
| *Avatime-familiar* | | | |
| A woman ties a baby onto her back |  | 0:48 | 47 |
| A woman prepares and sells a bowl of a local style of porridge |  | 0:51 | 51 |
| A woman heats water to bathe (she first cuts wood, lights a fire, then places water on the fire) |  | 6:04 | 364 |
| A man taps a palm tree to extract palm wine |  | 5:45 | 345 |

| *Dutch-familiar* | | | |
|---|---|---|---|
| A woman puts a bag on the back of her bicycle |  | 0:51 | 51 |
| A woman prepares a cheese sandwich |  | 1:38 | 98 |
| A woman changes the sheets on a bed |  | 5:30 | 330 |
| A man repairs a puncture in the inner tube of his bicycle |  | 5:58 | 357 |

All Dutch familiar events were filmed in the Netherlands with Dutch actors. They were selected according to the same criteria as the Avatime familiar events in consultation with three Dutch people, with the additional guideline that they should match the Avatime familiar events in terms of their length, the gender of the actor, and also be roughly comparable as discussed below.

The Avatime familiar event of a woman tying a baby on her back was matched with the Dutch familiar event of a woman putting a bag on her bicycle. They are both short events with durations of 48 and 51 seconds respectively. They both involve fixing something in place so it would not fall. They are both commonly

performed in public spaces, and thus are seen by all members of the community. Both events are also commonly performed by some, but not all, of the participants themselves: generally only girls will tie babies to their backs in Avatime communities and not all Dutch school children put their bags on the backs of their bicycles.

The Avatime familiar event of a woman selling porridge was matched with the Dutch event of a woman making a cheese sandwich. These are short events of 51 seconds and 1 minute 38 seconds, respectively. They are also both food preparation activities involving the use of multiple objects and are commonly seen by members of each community.

The Avatime familiar event of a woman heating water to bathe was matched with the Dutch familiar event of a woman making a bed. These are longer events taking 6 minutes 4 seconds and 5 minutes 30 seconds, respectively. They both make use of multiple objects and are clearly divided into distinct subparts when the actor steps out of the video frame to fetch new objects required for the task. They are both common domestic tasks either performed by the participants themselves or by their family members.

The Avatime familiar event of a man tapping a palm tree was matched with the Dutch familiar event of a man repairing a puncture in a bicycle inner tube. They are both complex tasks involving the use of multiple tools on a single object. These are also longer events of 5 minutes 45 seconds and 5 minutes 58 seconds, respectively. They are also both specialized tasks not commonly performed by the participants themselves, but which they have generally seen another person perform.

All events were filmed using a stationary high quality digital video camera positioned at head height[2]. The instructions to the actors were simply to perform the

[2] Some of the events were filmed outside and in public places. While care was taken to limit the amount of other activity within the video frame, three of the events – the woman tying a baby to her back, the woman selling porridge, and the woman heating water to bathe – showed other people or animals moving through the frame. These images where edited using Adobe Photoshop to remove signs of other activity. Unfortunately, due to the smoke from the fire and the number of animals moving around the courtyard it was not possible to remove all such other activity from the slideshow of the woman heating water to bathe. It was nevertheless clear

requested action e.g. heat water to bathe, or make the bed. The video was started when the actor indicated they were ready and stopped when they had finished, so the ultimate boundaries of the event were determined by the actor rather than the experimenter. As can be seen in Table 1, this freedom resulted in variation in the durations of the events. This variation has the side benefit for this experiment that participants cannot predict the end of a slideshow based on length alone.

In addition to the eight test items, there was one training item used to help familiarise participants with the task. This was a 34 second clip from the German cartoon 'Die Sendung mit der Maus'. In this clip a mouse brings a ball over to an elephant and they throw it back and forth a few times before the ball gets stuck inside the elephant's trunk. After a few attempts to get the ball out, the elephant blows it out and knocks the mouse over with the ball.

Slideshows of still images were formed for each of the videos, following the method of Hard et al. (2011): One still image was taken from the middle of each one second interval. All participants saw all events. The events from each culture were presented together in blocks and the order of the blocks was counter-balanced across participants. Within blocks, the events were shown in random order.

#### 5.3.1.3 Procedure

Participants were told they would see photos showing somebody doing something and they needed to press the space bar in order to see the next photo. Their task was to click through the photos so they understood what happened. They were told they would be asked questions once they had seen all of the photos. Instructions and responses were given verbally in the participant's native language.

Participants first saw the training slideshow of the mouse and the elephant playing ball, to become accustomed to clicking through the slideshows. The images were presented so the image filled the computer screen lengthways, the background above and below the image was black.

For each slideshow, participants first saw a slide with a single clause description of the goal of the action. For instance, the description for the training slideshow was "An elephant and a mouse play with a ball". Previous studies have shown that knowing the goal of an action has an influence on event segmentation (Wilder

---

that the main focus was on the woman and her activities, despite the occasional movement of chickens around the fireplace.

1978a; Zacks 2004). Thus, these descriptions were provided so that all participants knew the goal of the action but still differed in their familiarity with the event in general and the steps involved. The experimenter read the descriptions aloud for both Dutch and Avatime participants, since few of the Avatime participants were literate in Avatime.

Once the participant had clicked through all event slideshows, the experimenter showed them a single still image from each slideshow and asked them to describe what the person had done and to judge how familiar the action was to them. There is no good translation for "familiar" in Avatime, thus both Avatime and Dutch participants were asked whether they had ever performed the action themselves or seen somebody else do it, and if so how often. Their responses were scored on a scale from 1 to 5: 1 if they had never done the action or seen anyone else do it; 2 if they had done/seen it two or three times; 3 if they had done/seen it sometimes; 4 if they had done/seen it often; 5 if they had done/seen it very often. Participants were also asked to describe any differences between the way the event was performed in the slideshow and the way they had previously seen it performed.

The program Presentation was used to run the experiment and record the time between picture onset and when the participant pressed the button for the next picture. Verbal responses were audio recorded. Participants were tested individually and the experiment lasted between 20 and 30 minutes.

### 5.3.2 Results

The hypothesis was that events will be segmented more coarsely – with fewer event boundaries – by the people for whom they are familiar. This prediction was tested in two ways: using mixed-effects models and correlational analyses. The mixed-effects analysis (Section 5.3.2.3) provides a more direct test of the hypothesis but requires additional assumptions for calculating the number of event boundaries. The correlational analyses (Section 5.3.2.4) have the benefit of not requiring these additional assumptions, but are only able to determine whether the segmentation patterns of the two participant groups are similar or divergent. Before delving into each of these analyses, the participants' familiarity ratings are reported to confirm whether there was indeed a difference in familiarity between the two groups of stimuli (Section 5.3.2.1) and the calculation of the dwell time values from the raw

reaction times is described in Section 5.3.2.2. All analyses were performed using R 3.2.1 (R Core Team 2015), unless otherwise specified.

#### 5.3.2.1 Familiarity ratings

In order to check whether the events were in fact familiar or unfamiliar for each group, the familiarity ratings for Avatime and Dutch events were compared within each participant group. Since the ratings are non-parametric, this was done with Friedman tests in SPSS (IBM Corp 2011). As predicted, Dutch participants rated Dutch-familiar events as significantly more familiar (M = 3.74, SD = 0.44) than Avatime-familiar events (M = 1.04, SD = 0.09) $\chi^2(1) = 20.000$, $p < .001$. Conversely, Avatime participants rated Avatime-familiar events as significantly more familiar (M = 4.19, SD = 0.51) than Dutch-familiar events (M = 3.62, SD = 0.74) $\chi^2(1) = 12.250$, $p < .001$. While each group of participants rated events from their own cultural group as significantly more familiar than those of the other cultural group, there was a difference in how unfamiliar they found the events of the other group. The Dutch participants reported very little familiarity with the Avatime events, with an average rating of 1 or 'never seen or done previously'. In contrast, the Avatime participants rated the Dutch events with an average rating of 3 or 'seen or done sometimes'. Two of the Dutch-familiar events involved bicycles which had gone from being very rare to quite common in the Avatime region in the year between preparing the experiment and running it in the field. The other two events 'changing the sheets on a bed' and 'making a cheese sandwich' have similar correlates among the Avatime, who while they make their beds somewhat differently and don't use cheese still make their beds and eat buttered bread. The familiarity comparisons thus appear to be somewhat different between the two groups with the Dutch participants viewing familiar and highly unfamiliar events, while the Avatime participants view highly familiar and slightly familiar events.

#### 5.3.2.2 Calculating dwell times

The experimental procedure produces response times for each slide from each participant. These response times vary greatly between participants with some spending longer on each slide than others, for instance the baseline response time for each slide varied from 150 to 600ms. These raw response times also include artefacts common to other self-paced tasks. Following Hard et al. (2011), these raw response times were processed to produce "dwell times". This produces more

comparable values which are also more likely to reflect the participants' segmentation patterns.

Firstly, outlying data points were removed. These included two cases noted during testing where a participant paused to perform another action such as adjusting an item of clothing. Further outlying response times for individual slides were removed by inspection of the response time data for each participant for each event. Only clear outliers were removed. These were typically over 3 seconds and on average more than 11 – and no less than 5 – standard deviations from the mean. Twenty-three such outliers were removed. This is a tiny fraction of the more than 65,000 response times recorded.

As noted by Hard et al. (2011), participants tended to have longer response times for the first slide. This is probably because they are busy apprehending general aspects of the scene. The first slide from each event slideshow was therefore removed from all subsequent analyses.

As is common in reaction time studies, the response times had a positive skew (McCormack & Wright 1964). This means they were not normally distributed around the mean but instead had a larger tail on the right with more long rather than short response times. The skew was tested for each event individually using D'Agostino's skewness test with the moments package (Komsta & Novomestky 2012) in R and was significant ($p < .05$) for each event. The presence of a skew does not pose an essential problem for further analysis. However, since positive skews are a common feature of reaction time studies in general, it is likely a feature of the method rather than the segmentation patterns. The response times were therefore transformed with a $\log_{10}$ function, following Hard et al. (2011). This resulted in a smaller but still significant skew in the opposite direction.

There was a trend for participants to look longer at slides earlier in the slideshow and speed up as the slideshow progressed. As in Hard et al.'s (2011) study, this trend was well described by a power function. Power functions were thus individually fitted for each participant's viewing of each event. The goodness-of-fit of the power functions was tested by comparing the fit of the power function to the fit of a flat linear model using ANOVAs. The power function model provided a significantly better fit ($p < .05$) in 206 of the 320 cases (64.38%). The detrended residuals from these functions, referred to as the dwell times, were used in all subsequent analyses.

#### 5.3.2.3 Comparing the number of event boundaries

The number of event boundaries can only be estimated from the dwell times by making some assumptions. Firstly, according to Hard et al. (2011), event boundaries coincide with peaks in dwell times. So we can take peaks in the dwell time responses as indications of event boundaries. The prediction is then that Avatime participants should have fewer peaks for Avatime-familiar events and Dutch participants should have fewer peaks for Dutch-familiar events. That is, there ought to be an interaction between group and event type.

The influence of familiarity is also predicted to be stronger or in fact only apparent for coarse- rather than fine-grained event boundaries (Hard, Recchia & Tversky 2011; Zacks et al. 2007), so we need a way to distinguish between coarse- and fine-grained event boundaries. Hard et al. (2011) found the value of dwell times at peaks was related to the granularity level of the event boundary, with coarser event boundaries occurring with higher dwell time values than fine-level event boundaries. On average, Hard et al. (2011) found fine-level event boundaries occurred at slides with dwell times of 0.11 and coarse-level event boundaries occurred at slides with dwell times of 0.35. These average values are very similar to the average values found in another dwell time study (Meyer, Baldwin & Sage 2011). They were, therefore, used here as a guide for identifying event boundaries of different granularities.

Since the range in dwell time values varies across participants and events, these average values should not be used directly as absolute thresholds but rather adapted for each participant and event. This can be done via the standard deviation. The average standard deviation over all participants and events in the present study was 0.12. One standard deviation above the mean – which is always zero due to the way the residuals are calculated – is, thus, very close to the reported average for fine-level event boundaries. Three standard deviations above the mean is also close to the value reported for coarse-level event boundaries. Thus, one and three standard deviations above the mean can be used as reasonable thresholds for estimating the number of fine and coarse level event boundaries, respectively. They are sensitive to individual variation and on average correspond to previously reported values. As an example, the dwell times of one participant in response to the tying a baby on the back event is shown in Figure 5.2 with the thresholds at one and three standard deviations marked. In this case, there are four peaks above one standard deviation and one peak above three standard deviations. The peak at A occurs when the

woman finishes wrapping the cloth around her waist in preparation for tying the child. The peak at B corresponds with the woman bending down to pick up the child. The peak at C corresponds to the beginning of the actual tying action, as she places the cloth over the child's back. The final peak at D corresponds to the end of the tying, when she ties the final knot in the cloth. This final end of action peak is the only one which is above the threshold for coarse-level boundaries.

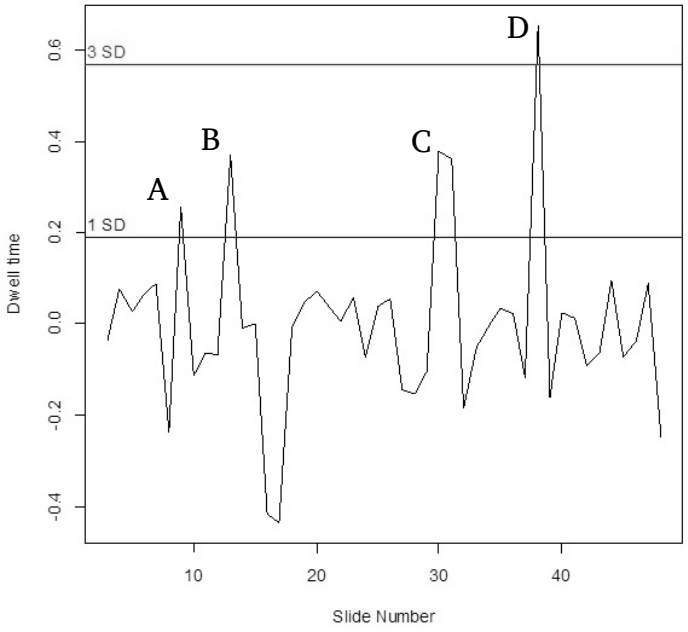



Figure 5.2: Dwell times for each slide of the woman tying a baby to her back slideshow from one of the Avatime participants (A17), with the thresholds at one and three standard deviations marked.

The number of peaks above each threshold was analysed using linear mixed effects with the lme4 package (Bates, Maechler, et al. 2015). This method of analysis has the benefit of testing the influence of experimental conditions while simultaneously accounting for variation across stimulus items and participants. The dependent variable was the number of peaks and due to the variation in the lengths of the slideshows this was taken to be the number of peaks per hundred slides for each event slideshow. The initial model included the experimental conditions of Culture (Avatime or Dutch) and Event type (Avatime or Dutch-Familiar). The Length of the stimulus (either long or short) and the Order in which it appeared in the experiment were also included (to account for possible influences of fatigue or other changes in segmentation patterns throughout the experiment), as well as all

interactions between these and the experimental conditions. Following recommendations by Barr, Levy, Scheepers, and Tily (2013) for avoiding anti-conservative biases, the initial model also included random slopes for each effect, either by Participant or by Stimulus as appropriate. Random slope terms were removed – in order of smallest to largest variance – only if required for the model to converge. Fixed effect terms were removed in a step-wise fashion always taking the term with the highest *p*-value so long as it did not participate in significant higher-order terms or cause a near significant reduction in the model's goodness-of-fit ($p < .1$). P-values were then calculated for all remaining fixed effects by comparison with the normal distribution, which is one of the standard methods for studies with this sample size (Barr et al. 2013). Since the coarse-level boundaries are a subgroup of the fine-level boundaries, the analysis is being performed twice over the same set of data. The significance criteria were, therefore, adjusted using Bonferroni's correction and only *p*-values less than .025 are considered to be significant.

The predicted interaction between Cultural group and Event type was not significant and was removed as part of the model selection process for both fine-grained $\beta = 0.558$, $SE = 0.650$, $z = 0.86$, $p = .398$, and coarse-grained event boundaries $\beta = 0.396$, $SE = 0.261$, $z = -1.52$, $p = .119$. So there were no differences in dwell time peak rate according to event familiarity.

When looking at the fine-grained boundaries, the mixed-effects model reduced to include only one fixed effect: Length. Peaks were more frequent in shorter event slideshows ($M = 10.04$) than longer ones ($M = 8.69$) $\beta = 1.310$, $SE = 0.331$, $z = 3.96$, $p < .001$. There were no other significant effects.

At the coarse-grained level, the final model also reduced to include only one main effect but here it was Cultural group and it was not significant according to the adjusted threshold $\beta = 0.369$, $SE = 0.184$, $z = 2.010$, $p = .044$. Numerically, Dutch participants had more frequent peaks above the coarse-level boundary ($M = 1.05$) than Avatime participants ($M = 0.68$). There were no other significant effects.

So the hypothesis was not borne out. The only significant factor in the frequency of dwell time peaks was the overall length of the stimulus item and this was only significant when considering all peaks above the fine-grained boundary. In the next section, I examine the whole dwell time responses, rather then just the peaks, to determine whether there are any general differences in how the two groups respond to the Avatime- and Dutch-familiar events.

#### 5.3.2.4 Correlations between dwell times

Previous work on explicit segmentation has shown people's segmentation patterns correlate with each other significantly more than chance, although the absolute correlations are small with average correlation values of around 0.17 (e.g. Zacks, Swallow, et al. 2006). Correlations between participants' dwell time values were examined here in two ways: using factor analysis with principle components extraction over all participants, and pairwise, following the method of Zacks et al. (2006) [3]. Factor analysis provides a global picture of the agreement among participants and whether or not they group together into clusters according to familiarity (e.g. Boster & Johnson 1989; Majid, Boster & Bowerman 2008). The pairwise method facilitates comparison with previous work by Zacks et al. (2006). If the Avatime and Dutch participants segment the events differently, they should form separate clusters in a factor analysis. Conversely, if they segment events similarly, then they should cluster together. Familiarity with an event has also been shown to improve alignment between coarse and fine level segmentations made by the same participant at different times (Zacks, Tversky & Iyer 2001). This suggests correlations between dwell time values may be higher among participants for whom the event is familiar. So, while an examination of the correlations between participants' dwell time values does not directly answer the question of whether people segment events more coarsely when they are familiar, it will provide an indication of whether people segment events differently according to familiarity.

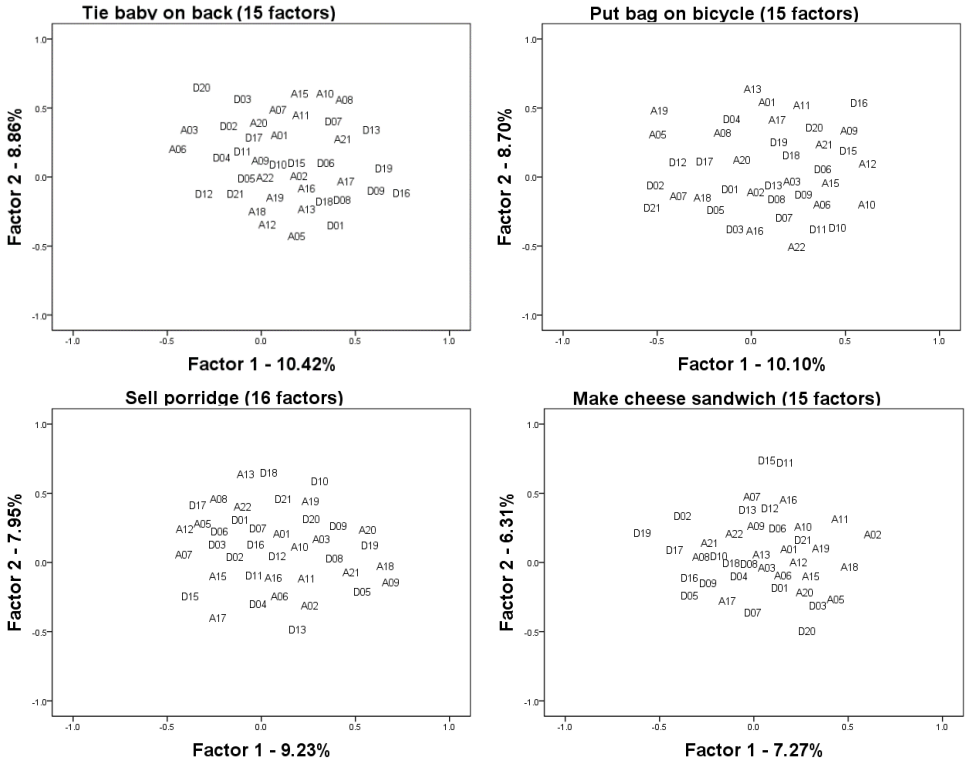
Factor analyses were conducted for all events using SPSS (IBM Corp 2011)[4], see Figure 5.3. The first component did not account for the majority of variance in any of the events. This suggests there is, in fact, little alignment between participants' dwell times. The Avatime and Dutch participants did not form distinct clusters, suggesting the differences in dwell time behaviour were no more different between the two participant groups than they were within them.

Correlations between dwell time values for all possible Avatime-Avatime, Dutch-Dutch, and Avatime-Dutch participant pairs were also calculated using Pearson's correlations for each event individually. The average correlation values for each

[3] I thank Jeff Zacks for recommending this analysis.

[4] Principle component analyses were also conducted for each event. The results were similar and only the factor analyses are reported here.

type of participant pairing and each event were then compared with chance using bootstrapped confidence intervals (Mooney & Duval 1993). The results were very similar across groups. The correlations for Avatime-Avatime participant pairs over all events ranged from -0.43 to 0.56, with an average of 0.01. For Dutch-Dutch pairs the correlations ranged from -0.46 to 0.51, with an average of 0.02. Looking between cultures, at the Avatime-Dutch pairs, the correlations ranged from -0.52 to 0.51 with an average of 0.01. The average correlations were not significantly different from chance for any group, for any event. A small percentage of individual participant pairs (10-30%, depending on condition) showed significant positive correlations. However, there were no differences according to group or event familiarity and these rates of significant correlations fall within the range one would expect to occur by chance when comparing this many participant pairs. This lack of correlation is in contrast to previous studies such as Zacks et al. (2006) where average correlations ranged from 0.05 to 0.19 according to experimental condition. This analysis, thus, verifies the factor analysis and confirms the large degree of variation among the individual dwell time patterns.



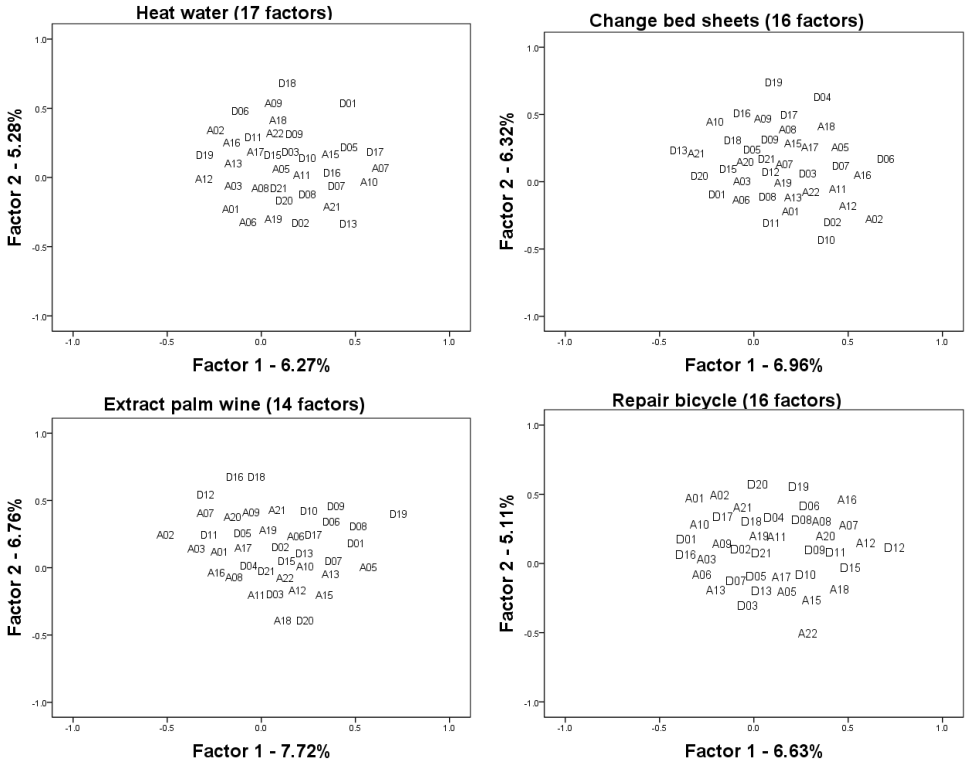



Figure 5.3: Loading plots for the factor analysis of each event, showing number of factors with eigenvalues greater than 1 (in brackets) and the percentage of variance accounted for by the first two components. Avatime and Dutch participants are indicated with an A or D, respectively, followed by their participant number.

#### 5.3.2.5 Comparison of dwell times to physical changes in the events

As a final test, dwell time values were compared to physical changes in the stimulus events. If people are segmenting events based on physical change, their dwell time values should, on average, correlate with those physical changes. In contrast, a lack of correlation would either suggest an alternative basis for the segmentation pattern or provide further corroboration of the high degree of individual variation.

Hard et al. (2011) tested this correspondence with physical change measured by the amount of pixel changes between consecutive images (after the images had been passed through a graphical filter to reduce variations in colour and lighting and focus on the edges of the actor and objects). They found 20 of the 30 participants showed reliable correlations between their dwell times and the amount of pixel change between slides. On average, the correlation between dwell times and amount of pixel change was 0.19 which was small but significantly different from zero. The same test was performed here with the prediction that most participants

would show a similarly small but significant correlation between dwell time values and amount of pixel changes.

There are several edge detection filters each of which is better suited to different types of images (Maini & Aggarwal 2009). Hard et al. (2011) were not specific regarding the filter they used. For the stimuli items in this study, the Sobel filter with a threshold of 50 was found to be the most appropriate, as it is less sensitive to noise and was better able to identify edges in the images. An example of one of the still images from the porridge selling event slideshow when processed with the Sobel filter can be seen in Figure 5.4. Once all slides had been passed through this filter, each pixel was paired with the corresponding pixel in the following slide and the absolute value of the difference in their brightness values was calculated. The total amount of pixel changes between two slides is then the sum of these differences for every pair of pixels.

The correlation of pixel change values with each participant's dwell time values was calculated using Pearson's correlations and bootstrapped confidence intervals (Mooney & Duval 1993) were used to evaluate the significance. The results varied considerably between slideshows, as can be seen in Table 5.2. Among the Avatime-familiar events, participants' dwell time values tended to correlate with pixel changes to a similar extent as in Hard et al. (2011). On average, Avatime participants showed significant positive correlations in all of the Avatime-familiar events, except for heating the water. Dutch participants' dwell times were also significantly correlated with the pixel changes, for the selling porridge and extracting palm wine events, but not for the heating water or tying a baby on the back events.

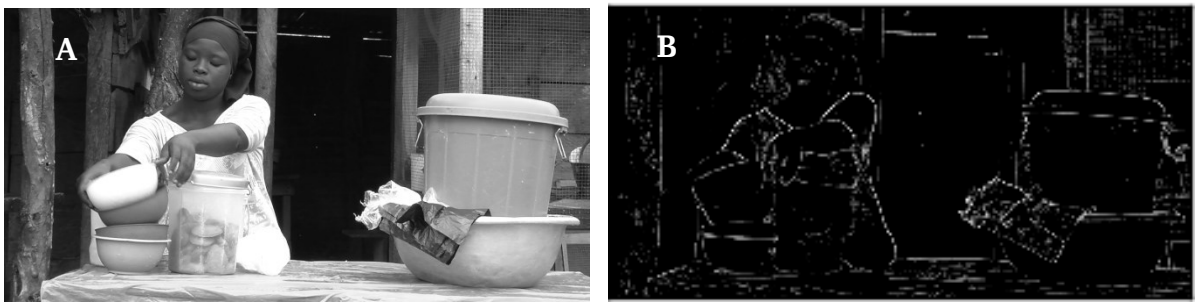



Figure 5.4: An image from the 'woman selling porridge' slideshow in A grey-scale and B after processing with the Sobel edge detection filter with a threshold of 50.

Table 5.2: The mean (M) correlations of dwell time values and standard deviations (SD) from each participant with pixel change values, for each slideshow. Mean correlations significantly different from chance are marked with an asterix (*).

| **Event Slideshow** | **Avatime participants** | | **Dutch participants** | |
|---|---|---|---|---|
| | **M** | **SD** | **M** | **SD** |
| *Avatime-familiar* | | | | |
| Tie baby on back | .20* | .22 | .06 | .20 |
| Sell porridge | .26* | .18 | .18* | .20 |
| Heat water | -.03 | .05 | -.01 | .06 |
| Extract palm wine | .15* | .17 | .18* | .17 |
| *Dutch-familiar* | | | | |
| Put bag on bicycle | .04 | .18 | .02 | .16 |
| Make cheese sandwich | .01 | .11 | -.04 | .11 |
| Change bed sheets | .02 | .07 | .06 | .06 |
| Repair bicycle | -.18* | .19 | -.11* | .18 |

Among the Dutch-familiar events, the correlations were generally no different from zero. The only event with average correlations significantly different from zero was the repairing the bicycle event and here both Avatime and Dutch participants showed significant negative correlations meaning their dwell times decreased when there when pixel changes increased.

The split between the Avatime- and Dutch-familiar events is intriguing. There is, however, no clear explanation for it. Some differences in the correlation rates between event items are likely due to differences in how well the pixel changes reflect the perceptual properties of the event. For instance, the movement of smoke during the heating water slideshow and the low light levels in the changing the bed sheets slideshow likely influenced the pattern of pixel changes leading them to deviate from the more relevant perceptual changes in these events. There is also a general difficulty in correctly identifying edges in image analysis (Radke et al. 2005), which could have led to variations across the stimulus items in how well the pixel changes reflect the relevant perceptual changes. Across all events, there was an impressive variation in correlation values suggesting some caution should be taken in interpreting the results. In particular, it will be informative to see whether the same events show similar correlation values in subsequent experiments.

### 5.3.3 Discussion

The prediction that familiarity influences the way events are segmented was not borne out. There were no systematic differences according to familiarity either when looking at the dwell time peaks or the dwell time responses taken as a whole.

The analysis of the numbers of dwell time peaks showed a difference at the fine-level of segmentation according to the length of the event. This has at least two possible explanations. There could be inherent differences in the event structures of the longer versus shorter events. For instance, it is likely event boundaries are more condensed in shorter events as compared to longer ones. The difference may also be due to differences in the level of attention participants maintain for different length items. A reduction in attention for longer items could manifest in less variation in dwell time values, and so less peaks. The restriction of this influence to the fine-level is consistent with previous findings suggesting fine-level segmentation is more influenced by perceptual properties of events, while higher-order event knowledge plays more of a role in coarse-level segmentation (Hard, Recchia & Tversky 2011; Zacks et al. 2007).

The correlation analyses showed high levels of variation in the individual participant's dwell times. This is in contrast to previous explicit studies of event segmentation which have found more consistency between participants (e.g. Hard, Tversky & Lang 2006; Newtson 1973; Newtson & Engquist 1976; Zacks, Tversky & Iyer 2001). This begs the question as to whether this consistency is in fact a product of the explicit segmentation tasks used. This question will be taken up further in the General Discussion (Section 5.6). The high degree of variation among participants may have been a limiting factor in the attempt to test the influence of familiarity. In the next experiment, participants are asked to describe each item immediately after viewing. Previous work (Cohen & Ebbesen 1979; Zacks, Tversky & Iyer 2001) suggests describing an event causes one particular event schema to be more strongly activated when forming a model of the current event. Thus describing the event is likely to increase any influence of familiarity as those who are familiar with it will have more well-defined event schemata. It is also likely to increase the correlations between participants by focusing their event models more sharply on a more limited range of event schemata.

## 5.4 Experiment 2 – Describing the Actor or Action

Previous work has shown greater agreement between fine- and coarse-grained segmentation of events when participants describe the events while segmenting them (Zacks, Tversky & Iyer 2001). The explanation given for this is that an event schema becomes more activated when describing an event and so exerts a stronger and more consistent influence over the segmentation. Asking people to describe events is thus likely to lead them to segment more uniformly and agree more with each other. Since a simultaneous description task would very likely disturb the implicit dwell time measure, participants were here asked to give their descriptions immediately after viewing each event slideshow. Cohen and Ebbesen (1979) have also suggested that different event schemata are activated depending on the particular description task. As in Cohen and Ebbesen's (1979) study, the present participants were asked to describe either the actor or the action. This offers an opportunity to test whether their finding that people describing actors segment events more coarsely than those describing actions extends to a different set of stimuli and an implicit measure of event segmentation. It is also quite possible that the type of description task will interact with familiarity, with participants being more influenced by the familiarity of the event when describing the actions.

### 5.4.1 Method

#### 5.4.1.1 Participants

Two new sets of participants were recruited for this experiment, none of which had participated in Experiment 1. Forty-nine Avatime participants, aged 10-18 (mean 14.1), were recruited at Junior High Schools in Amedzofe, Biakpa and Vane. Nine participants were tested but excluded from further analysis: three described the action when they were asked to describe the actor, three were interrupted during the task, two did not provide accurate descriptions of the stimuli, and one was not fluent in Avatime. All except for two of the Avatime participants had lived entirely within the Avatime traditional area. These two participants had been born in neighbouring regions and maintained close connections to the Avatime community since a young age. All Avatime participants also spoke English, all but one additionally spoke Ewe, four additionally spoke Twi, one additionally spoke Ga, and one also spoke French.

Forty-seven Dutch participants, aged 12-16 (mean 13.5), were recruited through advertisements in Nijmegen. Three participants were tested but excluded from further analysis: one due to technical problems, one for describing the action when they were asked to describe the actor, and one for not providing accurate descriptions of the stimuli. All of the Dutch participants had lived entirely within the Netherlands and were fluent in Dutch. Thirty reported at least some knowledge of English, thirteen some knowledge of French, ten some knowledge of German, three Spanish, one Indonesian, and one some knowledge of a language of the Maluku province of Indonesia.

#### 5.4.1.2 Materials

The same eight event slideshows and training slideshow from Experiment 1 were used in this study.

#### 5.4.1.3 Design

A between-participants design was employed with participants randomly assigned to one of two description conditions: actor or action. All participants saw all events: four were familiar to Avatime participants and four familiar to Dutch participants. As in Experiment 1, the familiar events from each culture were presented together in blocks and the order of the blocks was counter-balanced across participants. Within blocks, the events were shown in random orders.

#### 5.4.1.4 Procedure

The procedure was as in Experiment 1, with one exception. Instead of being told they would be asked some questions after completing viewing all slideshows, participants were asked to describe either the actor or the action – depending on condition – immediately following each slideshow. Once they had viewed and described all events, they were again shown a still image from each slideshow and asked to judge familiarity. The experiment lasted between 20 and 30 minutes.

### 5.4.2 Results

#### 5.4.2.1 Familiarity ratings

As predicted, Dutch participants again rated Dutch-familiar events as significantly more familiar (M = 3.78, SD = 0.46) than Avatime-familiar events (M = 1.10, SD = 0.21) $\chi^2(1) = 44.000$, $p < .001$. Conversely, Avatime participants rated

Avatime-familiar events as significantly more familiar (M = 4.11, SD = 0.58) than Dutch-familiar events (M = 3.45, SD = 0.84) $\chi^2(1) = 18.939$, $p < .001$.

#### 5.4.2.2 Calculating dwell times

Nineteen interruptions were noted while running the experiment and these points were from further analysis. Another twenty-two clear outliers with response times of more than 3 seconds were also removed. As in Experiment 1, the first slide from each event was excluded from further analysis. The response times were again significantly positively skewed for each event and a $\log_{10}$ transform was used which again resulted in a smaller but significant negative skew. Power functions were individually fitted for each participant's viewing of each event and ANOVAs were used to check their goodness-of-fit. The power function model provided a significantly better fit than the flat linear model in 517 of the 672 cases (76.93%). All further analyses were performed with the residuals from these power functions.

#### 5.4.2.3 Comparing the number of event boundaries

As in Experiment 1, the average standard deviation in dwell time values across all conditions was 0.12. Thus the number of fine-, and coarse-level event boundaries were again approximated by the number of dwell time peaks above one and three standard deviations from the mean, respectively.

The dependent variable in the linear mixed effects model was the number of peaks above each threshold per hundred slides. The initial model included the fixed effects Cultural group (Avatime or Dutch), Description condition (Actor or Action), Event type (Avatime- or Dutch-familiar), item Length (Long or Short), and Order, as well as all interactions between them. There were also random intercepts for Participant and Stimulus with random slopes for Cultural group, Description condition, Event type, Length, and Order. The random slope for Order by Participants had to be removed in all cases for the model to reach convergence. Since the analysis is performed over two subgroups, the significance threshold is adjusted to $p < .025$.

The primary prediction was that people would produce more dwell time peaks when the events were unfamiliar, so we should see an interaction between Cultural group and Event type. This influence of familiarity is predicted to occur particularly at coarser-level segmentation. It is also likely to be stronger for participants who are

describing the action rather than the actor since they should be more focussed on details of the event structure. Indeed, there was a significant interaction between Cultural group and Event type at the coarse-level of segmentation. However, at both the fine- and coarse-level Cultural group and Event type were involved in higher order interactions.

At the fine-grained level, the final model included the four-way interaction between Cultural group, Description condition, Event type and Length and a two-way interaction between Event type and Order as well as all lower order terms. The frequency of peaks above the fine-grained threshold tended to increase throughout the experiment for Avatime-familiar events, but decrease for Dutch-familiar events $\beta = -0.350$, $SE = 0.138$, $z = -2.539$, $p = .011$. The four-way interaction was also significant $\beta = -6.016$, $SE = 2.344$, $z = -2.566$, $p = .010$. Among the Long items, there was little variation and a post-hoc analysis showed no significant differences. Among the Short items, there was a split according to whether people were describing the Action or Actor. The people describing the Action showed the predicted influence of familiarity in response to the Dutch-familiar events only: Avatime participants had more peaks ($M = 10.35$) than Dutch participants ($M = 8.30$). There was no difference between Avatime ($M = 7.62$) and Dutch participants describing the Action ($M = 7.77$) in response to the Avatime-familiar events. When people described the Actor, they showed an influence of familiarity in the opposite direction. Avatime participants had more peaks ($M = 9.30$) than Dutch participants ($M = 7.92$) for Avatime-familiar events, and Dutch participants had more peaks ($M = 10.11$) than Avatime participants ($M = 8.89$) for Dutch-familiar events. There were no other significant effects. In particular, there was no main effect of Description and thus no evidence for those describing the action to segment more finely than those describing the actor as Cohen and Ebbesen (1979) found in their explicit segmentation study.

At the coarse-grained level, the final model was made up of a three-way interaction between Cultural group, Event type, and Length, as well as a main effect of Order. There was a trend for the peak frequency to increase throughout the experiments, but it did not reach significance given the adjusted threshold $\beta = 0.039$, $SE = 0.018$, $z = 2.143$, $p = .032$. Dutch participants had more frequent peaks overall than Avatime participants $\beta = 0.669$, $SE = 0.203$, $z = 3.292$, $p = .001$. Cultural group was also involved in two separate two-way interactions one with Length and the other with Event type. Avatime participants had more frequent peaks for Short

items, whereas Dutch participants had slightly more frequent peaks for Long items β = -1.391, SE = 0.286, $z$ = -4.871, $p$ < .001. The hypothesis concerned the interaction between Cultural group and Event type. However, while this interaction was significant, the pattern of results was not as predicted: Dutch-participants had more frequent peaks for Dutch-familiar events than Avatime participants did and there was little difference between the two groups in response to the Avatime-familiar events β = -0.835, SE = 0.291, $z$ = -2.858, $p$ = .004. There was also a significant higher order interaction between Cultural group, Event type and Length β = 1.872, SE = 0.392, $z$ = 4.780, $p$ < .001, see Figure 5.5. For Long items, the difference according to familiarity was as predicted: Avatime participants had more frequent peaks (M = 0.64) than Dutch participants (M = 0.48) for Dutch-familiar events, whereas Dutch participants had more frequent peaks (M = 1.54) than the Avatime participants (M = 0.84) for Avatime-familiar events. Among the Short items, however, this pattern was reversed and Dutch participants had more frequent peaks (M = 1.23) than the Avatime (M = 0.91) for Dutch-familiar events, while Avatime participants had more frequent peaks (M = 1.37) than the Dutch (M = 0.64) for Avatime-familiar events. There were no other significant effects. So again there was no evidence of a difference according to description task.

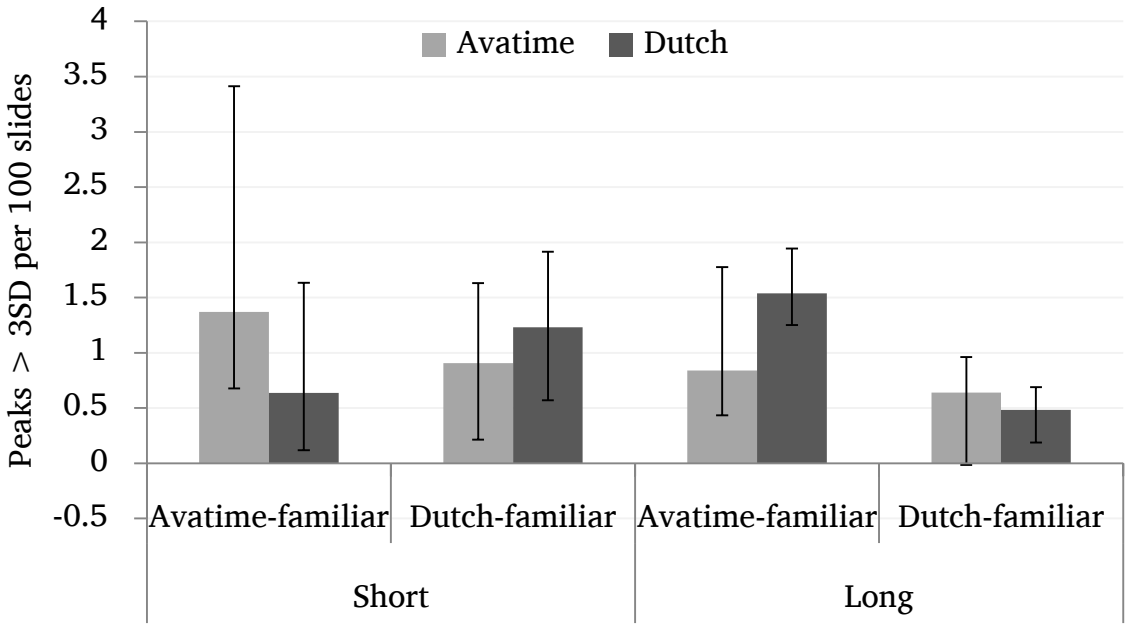



Figure 5.5: Average peaks above three standard deviations per 100 slides for each type of stimulus item (Long or Short, and Avatime- or Dutch-familiar) by each group of participants. Error bars indicate 95% confidence intervals.

Overall then, there was an influence of familiarity on the frequency of dwell time peaks, but whether it related to coarser or finer segmentation appeared to depend on the length of the stimulus and whether the person was focused on describing the actor or the action.

#### 5.4.2.4 Correlations between dwell times

A factor analysis with principle component extraction was conducted over the dwell time values for each event to determine whether people were segmenting the events following universal or distinct strategies. If everyone segmented the events similarly, the majority of the variance should be accounted for by a single factor for which the participants are positively loaded. In contrast, if Avatime and Dutch people employ distinct strategies, the majority of variance should still be accounted for by a single factor but the two participants should form separate clusters. Finally, the participants may show no shared segmentation style either across or within the cultural groups. In this case the majority of the variance would not be accounted for by a single factor. As in Experiment 1, the factor analyses show a large degree of individual variation. The first component does not account for a majority of the variance and there are no distinct clusters, see Figure 5.6. This suggests the participants do not have a common pattern to their dwell time values either universally or by condition.

Looking at the correlations between pairs of participants, the correlations over all events for pairs of Avatime participants describing the actor ranged from -0.44 to 0.62 ($M = 0.01$) and those describing the action ranged from -0.45 to 0.47 ($M = 0.01$). For pairs of Dutch participants the correlations ranged from -0.44 to 0.63 ($M = 0.02$) when describing the actor and from -0.42 to 0.53 ($M = 0.04$) when describing the action. Only 18-27% of individual participant pairs showed a significant positive correlation and there were no differences according to the type of participant pair, the description condition, or whether or not the event was familiar. This corroborates the factor analysis and Experiment 1 in confirming the lack of agreement in dwell time patterns.

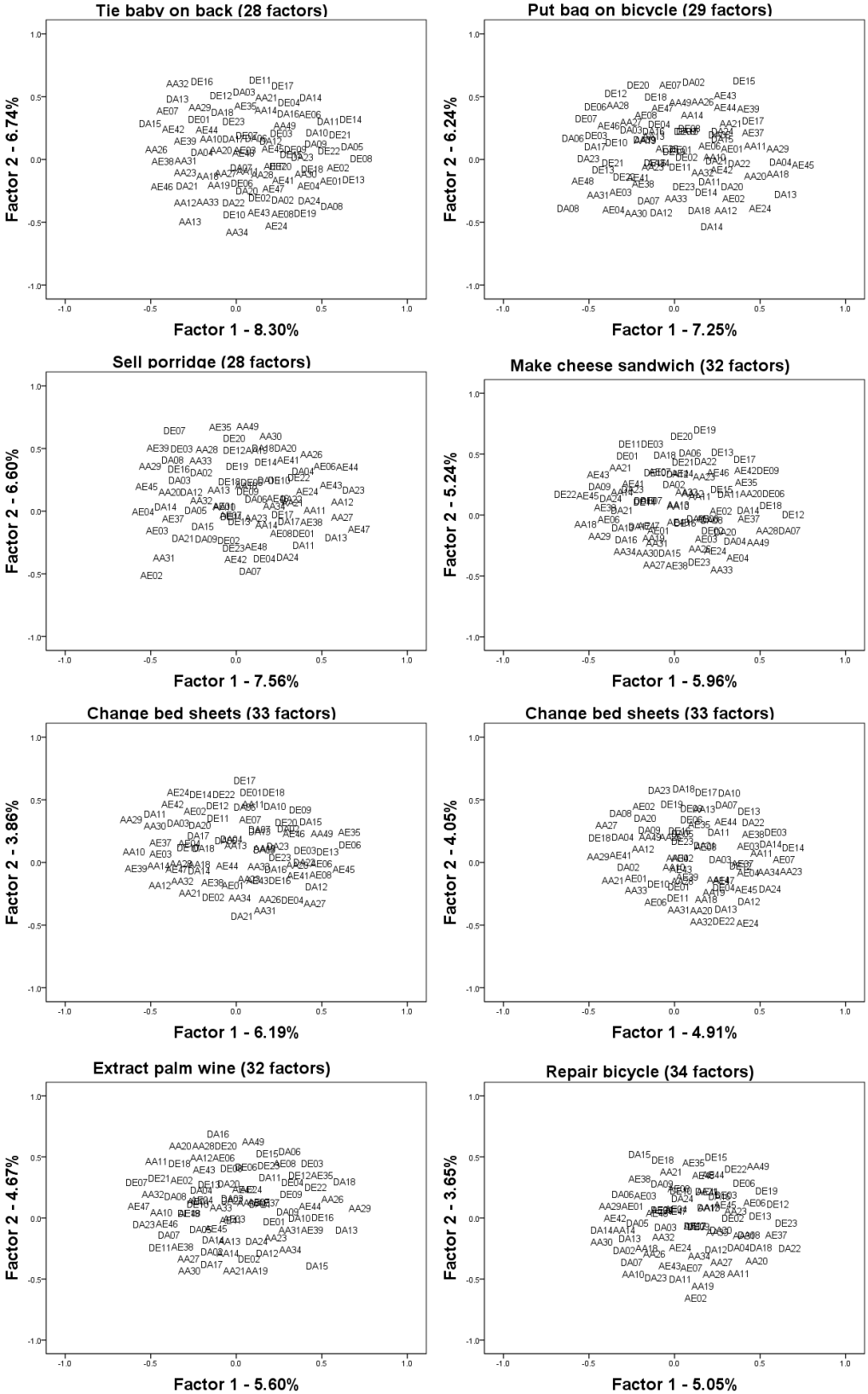



Figure 5.6: Loading plots for each event, showing the number of factors with eigenvalues greater than 1 and the percentage of variance accounted for by the first two components. An initial D or A indicates Dutch or Avatime participant,

respectively. The following A or E indicates whether the participant was describing the Actor or Event, respectively. This is followed by the participant number.

#### 5.4.2.1 Comparison of dwell times to physical changes in the events

As in Experiment 1, each participant's dwell times for each slideshow were compared to the pixel changes between slides. The results are shown in Table 5.3. Unlike Experiment 1, the correlations were uniformly low and not significantly different from zero. This calls into question the reliability of the correlations found for the Avatime-familiar events in Experiment 1.

Table 5.3: This table shows the mean (M) and standard deviation (SD) of correlations of the dwell time values from each participant with the pixel change values, for each slideshow.

| | **Avatime participants** | | | | **Dutch participants** | | | |
|---|---|---|---|---|---|---|---|---|
| | **Actor** | | **Action** | | **Actor** | | **Action** | |
| **Event Slideshow** | **M** | **SD** | **M** | **SD** | **M** | **SD** | **M** | **SD** |
| *Avatime-familiar* | | | | | | | | |
| Tie baby on back | -.02 | .13 | .01 | .16 | -.02 | .15 | -.01 | .16 |
| Sell porridge | .05 | .14 | .02 | .12 | .04 | .13 | .07 | .13 |
| Heat water | -.01 | .08 | -.02 | .08 | .02 | .09 | .01 | .08 |
| Extract palm wine | -.02 | .13 | -.01 | .10 | .03 | .14 | .02 | .10 |
| *Dutch-familiar* | | | | | | | | |
| Put bag on bicycle | .01 | .21 | .00 | .18 | -.01 | .18 | -.16 | .14 |
| Make cheese sandwich | .02 | .07 | .02 | .07 | .03 | .14 | .05 | .09 |
| Change bed sheets | .01 | .05 | .00 | .06 | .02 | .05 | .02 | .07 |
| Repair bicycle | -.02 | .13 | -.01 | .12 | .02 | .08 | -.03 | .09 |

### 5.4.3 Discussion

There were three predictions for this experiment. The first was that the explicit description task would lead to more agreement in dwell time patterns between participants. The second was that people would segment events more finely when asked to describe actions as opposed to actors (c.f. Cohen & Ebbesen 1979). The third was that people segment events more coarsely when they are familiar. None of these predictions were borne out as expected.

The correlation analyses showed a similar level of individual variation as in Experiment 1 with correlations uniformly low and not different from chance. The comparison with the pixel changes showed little correspondence, with no events showing robust correlations between participants' dwell time values and pixel change. So the description task did not lead to an overall increase in agreement between participants.

There was no evidence of the general influence of description task reported by Cohen and Ebbesen (1979). This may be a difference between explicit and implicit segmentation tasks.

Finally, while there was an influence of familiarity, the nature of that influence differed for short and long events. For long events, the results were as predicted with people who were familiar with the event segmenting it more coarsely. However, the pattern was reversed for short events, with familiarity coinciding with finer-segmentation. This pattern was clearer at the coarse-level of segmentation. At the fine-level, there was also a difference according to description condition, and the differences in response to long events were not as clear. There are several possible explanations for the split between long and short events. It could be that the short events in this experiment were simply too short and this limited the analysis, or it could be that familiarity influences events of different lengths in different ways. These possibilities are discussed further in the General Discussion (Section 5.6).

## 5.5 Experiment 3 – Priming with SVCs and Coordinated clauses

The previous two experiments focused on whether or not culturally familiar events are segmented more coarsely than unfamiliar events. The other question raised in this chapter was whether the use of serial verb constructions (SVCs) influences event segmentation independently of familiarity. The present experiment provided an initial foray into investigating this possibility by priming participants with SVCs or coordinate clauses before they viewed the event slideshows. Participants primed with SVCs were predicted to segment more coarsely, i.e. with fewer event boundaries, than those primed with coordinate clauses.

### 5.5.1 Method

#### 5.5.1.1 Participants

Fifty Avatime participants, aged 12-18 (mean 14.4), were recruited at Junior High Schools in Amedzofe, Gbadzeme and Vane. None of them had participated in either Experiment 1 or 2. Ten participants were tested but excluded: four due to difficulty priming, three due to external interruptions, two due to inattention, and one due to technical problems. All except two of the participants had lived entirely within the Avatime traditional area, and these two had maintained close connections with the Avatime area since a young age. All participants were fluent in Avatime. They also all spoke English and Ewe, eight additionally spoke Twi, two additionally spoke Ga, one Hausa, one Yoruba, and one Nyagbo.

#### 5.5.1.2 Materials

The same event slideshows from Experiment 1 were also used in this experiment. An additional set of pictures and sentences were used for priming as described within Section 5.5.1.4 below.

#### 5.5.1.3 Design

A between-participants design was used with participants randomly assigned to either the SVC or Coordinate priming condition. Each participant saw all the Avatime- and Dutch-familiar events. Since there are only Avatime participants, these events will simply be referred to as familiar or unfamiliar, respectively. The familiar and unfamiliar events were presented together in blocks and the order of the blocks was counter-balanced across participants. Within blocks, event slideshows were shown in random orders.

#### 5.5.1.4 Procedure

Participants first heard the instructions, in Avatime, for the slideshow task as in Experiment 1, and clicked through the training slideshow. They then performed a priming task designed to prime them with either the SVC or coordinate clause construction, before proceeding with the test items.

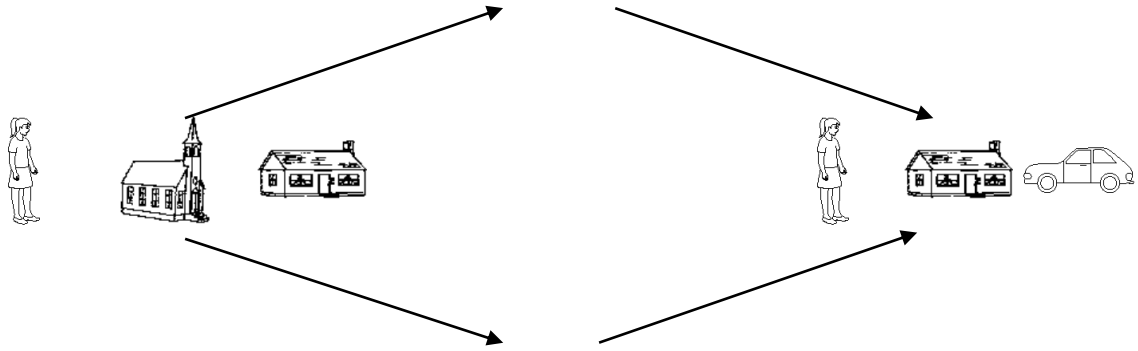
The priming task was adapted from structural priming studies such as Bock (1986) and Bock and Griffin (2000). Participants saw a set of pictures depicting the arguments of an event (presented simultaneously) and heard a recorded sentence describing the event played through a set of headphones. The recorded sentences

were either SVCs or two coordinated clauses depending on condition. Participants repeated the sentence they heard to the experimenter, who had not heard it. They then saw a similar set of pictures depicting the arguments of another event and were asked to describe it. This sequence of repeating a sentence and then describing another set of pictures was repeated a total of 30 times with different sets of pictures depicting different events. An example is shown in the diagram in Figure 5.7[5]. The first six repeat/describe pairs related to similar events from the same event frame of someone moving from one place to another. The subsequent repeat/describe pairs related to different event types with the event frame changing after every third pair. Participants were very quick to grasp the task and could do it easily after the first couple of repeat/describe pairs.

The design of the priming task biased participants towards repeating the syntactic structure they had heard: A single syntactic structure – either SVC or coordinated clauses – was used for all items and the picture stimuli were diagrammatic, showing how to fill in the argument slots rather than depicting the event. This design was employed since the purpose was to make the prime as strong as possible to ensure participants used the primed construction, rather than the more typical goal of testing the extent of the priming effect (e.g. Bock 1986; Bock & Griffin 2000). All except four participants – who were excluded from subsequent testing – were successfully primed and consistently used the primed construction after the first few event frames.

[5] Most pictures were taken from the Snodgrass and Vanderwart (1980) set. Additional pictures with similar line styles were used where necessary such as for the boy and girl, which are not included in Snodgrass and Vanderwart. Although these pictures were not designed to match local Avatime norms, they were all readily recognizable by the Avatime participants. They were also always introduced with an Avatime label and only served as placeholders for these entities in the priming task and so any issues due to nontypicalness should be limited.

a. *ɔ́-dzɛ* *a-dɔ́* *sɔlɔmɛ*
$C_1$s-woman $C_1$s.PFV-move.from:LOC church
*ba* *ke-pe = a* *mɛ̀*
come $C_6$s-house = DEF inside
'The woman moved from the church to the house.'



b. *ɔ́-dzɛ* *a-dɔ́* *sɔlɔmɛ*
$C_1$s-woman $C_1$s.PFV-move.from church
*lɛ̌* *a-ba* *ke-pe = a* *mɛ̀*
and $C_1$s.PFV-come $C_6$s-house = DEF inside
'The woman moved from the church and came to the house.'

Figure 5.7: A diagram of one of the repeat and describe sequences. The three pictures on the right were shown to the participant while they listen to either the SVC in (a) or the coordinated clauses in (b) depending on which condition they were in. They repeated this description to the experimenter. They then saw the similar set of pictures on the right and were asked to describe it. In this case, they would produce descriptions like 'The woman moved from the house to the car' or 'The woman moved from the house and came to the car'.

After this initial priming task, participants continued with the event slideshows as per the procedure in Experiment 1. Before each slideshow, the prime was refreshed by repeating six repeat/describe priming pairs. In most cases, participants responded with the primed construction in five or six out of these describe trials. If they used another construction more than once, or they used another construction on the final trial, they did another six repeat/describe pairs. The order of the event frames used was consistent between participants, but since the order of event

slideshows was random, the pairings of priming events and slideshow events was also random.

After viewing all event slideshows, participants were shown a still image from each slideshow and asked to describe what the person had done and rate how familiar that activity was to them as per the procedure used in Experiment 1.

The program Presentation was used to run the experiment and record the time between displaying each picture and when the participant pressed the button for the next picture. Verbal responses were audio recorded. Participants were tested individually and the whole experiment lasted approximately 45 minutes.

### 5.5.2 Results

#### 5.5.2.1 Familiarity ratings

Avatime-familiar events were again rated significantly more familiar ($M = 4.22$, $SD = 0.58$) than Dutch-familiar events ($M = 3.35$, $SD = 0.93$), $\chi^2(1) = 19.882$, $p < .001$.

#### 5.5.2.2 Calculating dwell times

One interruption was noted during the experiment. This and another seven clear outliers with response times of more than 3 seconds were removed. As in the previous experiments, the first slide from each event was also excluded. There was again a significant positive skew, which was transformed with a $\log_{10}$ transformation to produce a smaller but still significant negative skew. The power function models provided a significantly better fit than the linear model in 244 of the 320 cases (76.25%).

#### 5.5.2.3 Comparing the number of event boundaries

The average standard deviation among the dwell times was 0.10. This was a little lower than in the previous studies. One standard deviation is, however, still very close to the previously reported average dwell time for fine-grained event boundaries of 0.11 (Hard, Recchia & Tversky 2011). Three standard deviations is somewhat lower than the previously reported average value of 0.35 (Hard, Recchia & Tversky 2011), but for consistency with the previous experiments it was again used here, although a threshold of 3.5 standard deviations would have been a reasonable alternative. The number of peaks above each threshold per hundred

slides was again analysed using linear mixed effects. The initial model included Priming condition (SVC or Coordinate), Event type (Familiar or Unfamiliar), item Length (Long or Short), and Order of the stimuli, as well as all interactions between them. Random intercepts for Participant and Stimulus were included with random slopes for Event type, Length and Order or Priming condition respectively. The threshold for significance was again adjusted to $p < .025$. The prediction is that participants primed with SVCs should have less frequent peaks than participants primed with coordinate constructions. Since SVCs are used when treating the event as a single unified whole (Chapter 4) and so their use may encourage a more unified discretization pattern. Familiarity is also predicted to influence event segmentation with participants having less frequent peaks with familiar events, or indeed, following Experiment 2, possibly more so for the short items.

Looking at the fine-level boundaries, the model reduced to include two-way interactions between Priming condition and Event type and between Event type and Order, as well as a main effect of Length. As predicted, participants primed with SVCs had fewer peaks (M = 9.29) than those primed with Coordinate constructions (M = 10.11) $\beta = -1.407$, SE = 0.515, $z = -2.732$, $p = .006$. It appears this difference may largely be driven by the Familiar events where the difference between SVC (M = 9.08) and coordinate priming conditions (M = 10.48) is much greater than among the Unfamiliar events (M = 9.49 vs. M = 9.75), see Figure 5.8. However, this two-way interaction did not reach significance with the adjusted significance threshold $\beta = 1.191$, SE = 0.607, $z = 1.962$, $p = .049$. Event type was also significant as a main effect: Familiar events had more frequent peaks above the fine-level boundary overall (M = 9.78) than Unfamiliar events (M = 9.62) $\beta = -2.884$, SE = 0.785, $z = -3.673$, $p < .001$. There was also a tendency for the frequency of peaks to decrease throughout the experiment for Familiar events but increase for Unfamiliar ones $\beta = 0.470$, SE = 0.147, $z = 3.194$, $p = .001$. Finally, as noted in the previous experiments, there was a tendency for shorter events to have more frequent peaks above the fine-level event boundary (M = 10.52) than longer events (M = 8.88) $\beta = 1.642$, SE = 0.364, $z = 4.510$, $p < .001$.

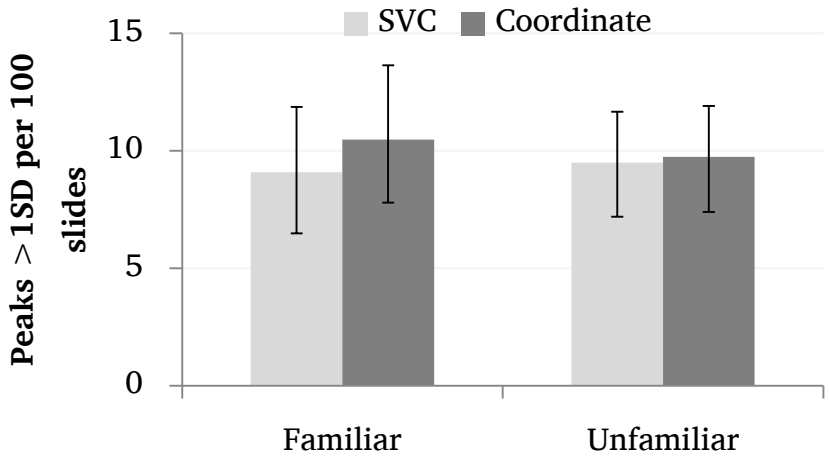



Figure 5.8 Peaks above one standard deviation per 100 slides for Familiar and Unfamiliar events by participants primed with either SVCs or Coordinate clauses. Error bars indicate the 95% confidence interval.

For the coarse-level boundaries, the final model included a three-way interaction between Priming condition, Length, and Order as well as all lower-order terms. However, none of these effects reached significance.

#### 5.5.2.4 Correlations between dwell times

As in the previous experiments, the factor analyses showed a large degree of individual variation: The first component did not account for a majority of the variance and there were no distinct clusters among the participants, see Figure 5.9.

This variation was again confirmed with pairwise correlations. The correlations between pairs primed with SVCs ranged between -0.49 and 0.43, as did those between pairs primed with coordinate clauses. Correlations of between-group pairs ranged from -0.46 to 0.57. In all cases the average correlation was 0.01 and not different from chance. Only 19-20% of individual participant pairs showed a significant positive correlation within each event and there were no differences according to priming condition or familiarity.

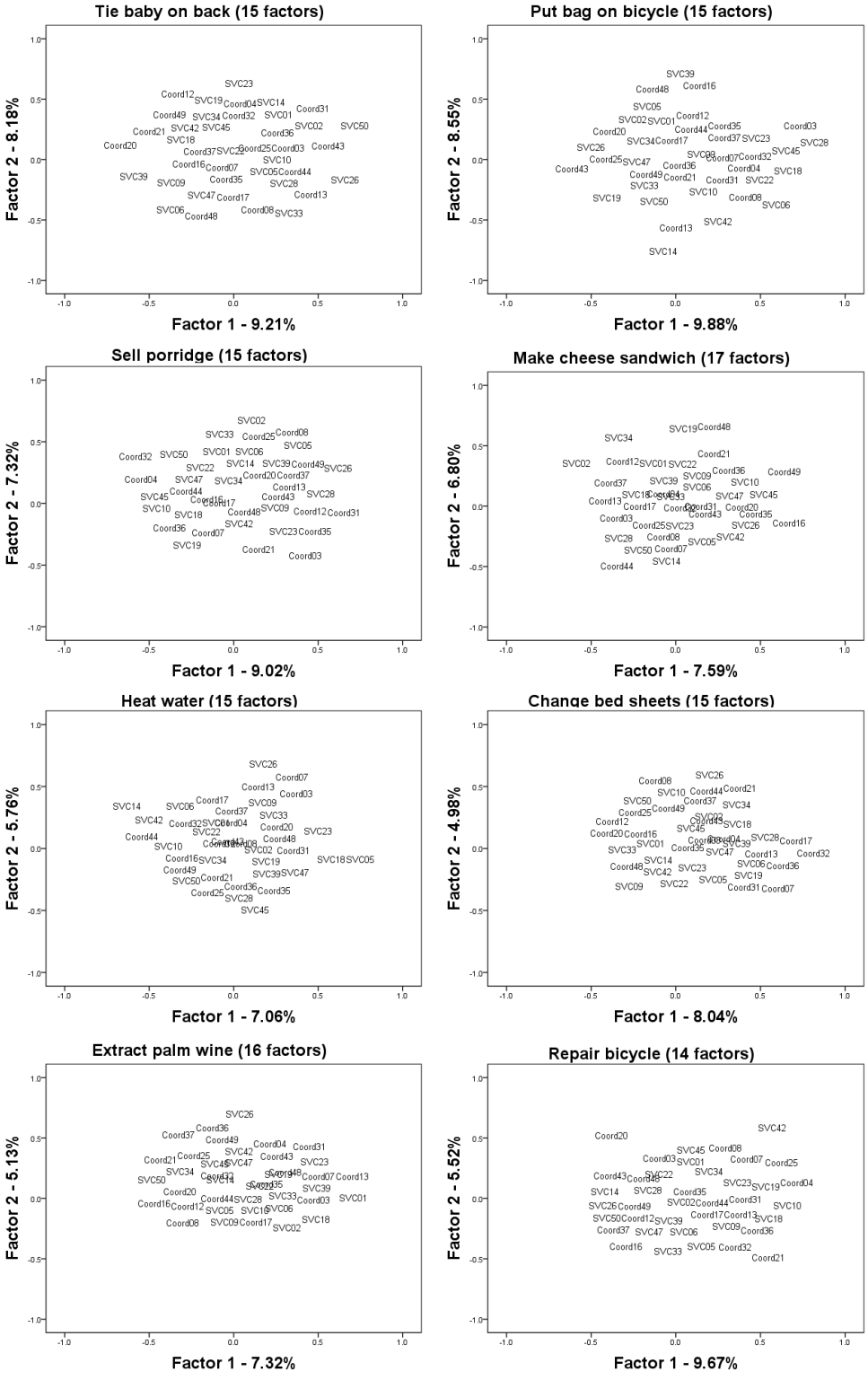



Figure 5.9: Loading plots for each event, showing the number of factors with eigenvalues greater than 1 and the percentage of variance accounted for by the first two components.

#### 5.5.2.1 Comparison of dwell times to pixel change

The average correlations of dwell times with pixel changes were again uniformly low and not significantly different from zero, see Table 5.4.

Table 5.4: Mean (M) and standard deviation (SD) of correlations of dwell time values with pixel change from each participant, for each slideshow.

| | **SVC** | | **Coordinate** | |
|---|---|---|---|---|
| **Event Slideshow** | **M** | **SD** | **M** | **SD** |
| *Familiar* | | | | |
| Tie baby on back | .00 | .13 | .03 | .10 |
| Sell porridge | -.01 | .12 | .04 | .14 |
| Heat water | -.02 | .07 | -.03 | .05 |
| Extract palm wine | .08 | .07 | .00 | .08 |
| *Unfamiliar* | | | | |
| Put bag on bicycle | .03 | .20 | .04 | .14 |
| Make cheese sandwich | .00 | .10 | .00 | .07 |
| Change bed sheets | .02 | .07 | -.03 | .08 |
| Repair bicycle | -.04 | .14 | -.03 | .13 |

### 5.5.3 Discussion

The goal of this experiment was to test whether the use of SVCs had an influence on event segmentation independent of familiarity. The prediction was that participants primed with SVCs would have fewer event boundaries – dwell time peaks – than those primed with coordinate clauses. This prediction was confirmed for fine-level event boundaries. Participants primed with SVCs appeared to group actions together more and so have fewer dwell time peaks than those primed with coordinate clauses. The fact that there appeared to be a larger difference in response to the Familiar events suggests the event schema is an integral part of the mechanism of this priming influence. Priming with SVCs likely triggers the activation of more holistic event schemata with sub-events grouped together while priming with coordinate clauses likely triggers activation of more granular schemata. When the event is unfamiliar, the event schemata are not well-developed and the different construction primes have less impact.

It is notable that this difference was only apparent when looking at all peaks above the fine-level threshold, despite suggestions that this level of segmentation is more perceptually driven than coarser-level units (Hard, Recchia & Tversky 2011; Zacks et al. 2007). One possible explanation for this is that while both SVCs and coordinate clauses have similar coarse-level event boundaries, coordinate clauses explicitly mark the finer-level boundary within this larger unit where SVCs minimise this fine level boundary and focus on the single coarse-level unit. So the difference between the two constructions in fact lies at the fine-level of segmentation.

## 5.6 General discussion

The present set of experiments aimed to test two potentially intertwined hypotheses: that people who are familiar with an event segment it more coarsely than those who are unfamiliar with it, and that the use of SVCs encourages coarser-grained segmentation than that of coordinate clauses.

Familiarity with an event was predicted to lead participants to segment it more coarsely (e.g. Hard, Tversky & Lang 2006; Newtson 1973; Wilder 1978b). An influence of familiarity was noted in the second experiment, though not the first. This difference between the experiments was likely due to the greater number of participants in the second experiment as well as the description task which likely increased activation of and attention to a particular event schema (Cohen & Ebbesen 1979; Zacks, Tversky & Iyer 2001). Unexpectedly, the influence of familiarity varied according to the length of the event. The predicted pattern occurred only among the longer stimulus items, whereas responses to the shorter events showed the reverse pattern, with familiarity coinciding with finer-grained segmentation. There are at least three potential explanations for this split between long and short events. The simplest explanation is a statistical one: the short events may not have provided a sufficient number of data points to yield reliable results. 60% of all viewings of the shorter events included no peaks above the coarse-grained boundary, compared to 16% of the longer events. The short events, particularly at the coarser-grains of segmentation, may thus have been limited by a floor effect where the results were skewed by the high number of zero values. However, the reasonably large effect sizes, suggest this is a real difference rather than an error due to limited data. Another possibility is that the short events may have been too short for unfamiliar viewers to form a reasonable event model and this may have limited the influence

of familiarity. The final possibility is that familiarity does not in fact lead to finer segmentation in general, but rather to a more default spacing of boundary points. When people are unfamiliar with an event they do not have a well-developed event schema for it, this could lead them to misjudge the global event structure and the relative importance of any potential event boundaries they perceive. This may lead them to break longer events into relatively fewer pieces but shorter events into more. This explanation also receives some support by the tendency for dwell time peaks to be condensed, occurring more frequently in the shorter events than in the longer ones. It is worth noting that previous studies have typically used longer stimuli events of 5 (Newtson 1973; Wilder 1978b), or 10 minutes (Zacks 2004). The stimulus used by Hard et al. (2006) comes the closest to the short items in the present study, though at 84 seconds it was still longer than many of the shorter items used here. Further testing with stimulus items of different lengths would help isolate the mechanisms behind these differences.

While the present studies showed an unexpected split in the influence of familiarity on long versus short action sequences, they do present some evidence for an influence of general familiarity. This is in line with previous studies which have shown influences of predictability (Wilder 1978b; Zacks 2004), newness (Hard, Tversky & Lang 2006), and unexpected interruptions t(Newtson 1973). The present finding that cultural familiarity with an event can influence the way a person segments it is predicted by Zacks et al.'s (2007) Event Segmentation Theory and the general claim that event segmentation is influenced by event schemata.

The other primary goal of this chapter was to test the influence of SVCs on event segmentation. Participants primed with SVCs indeed had less frequent dwell time peaks than those primed with coordinate clauses. This difference was found only above the fine-grained threshold, which appears to go against previous findings that finer-levels of segmentation are more influenced by the raw perceptual properties of events and less by top-down conceptual factors than coarser-levels. It is, however, in line with the nature of the difference between SVCs and coordinate clauses where the two have the same coarse level boundaries but coordinate clauses explicitly distinguish an additional finer-level boundary.

It was initially predicted that the influences of SVCs and familiarity could work together or confound each other by both encouraging coarser segmentations. There was no evidence of such an interaction. There was, however, a suggestion that

construction type may have more influence on familiar events. This would indicate that the primed construction influences segmentation via an influence on the event schemata participants activate for each event.

This initial study has thus documented the possibility of an influence of syntactic segmentation on event segmentation during perception and opened the door for further investigations. The positive result raises questions for future research, such as: Does the difference arise because SVCs encourage coarser segmentation, because coordinate clauses encourage finer segmentation, or both? What is the mechanism behind the influence? Do natural online language use or habitual description patterns also have an influence?

Finally, there was more variability between participants than expected. While individual variability has been noted in previous studies, there has nevertheless been significant agreement among participants (e.g. Cohen & Ebbesen 1979; Hard, Recchia & Tversky 2011; Zacks, Swallow, et al. 2006; Zacks, Tversky & Iyer 2001). In contrast, the dwell times of participants in the present studies did not tend to correlate with each other. This could be due to a difference between implicit and explicit segmentation tasks, the latter of which could well encourage more normative segmentation patterns. There have so far been very few studies working with the implicit dwell time measure and none have examined the individual variation so this study breaks new ground and supplies valuable missing data to the field. If the more normative segmentation patterns are indeed a product of the explicit segmentation task, it raises important questions about the generalizability of previous findings and the nature of the event segmentation process. This would need to be tested in future work comparing the two tasks in more detail.

## 5.7 Conclusion

The present chapter aimed to test the possible influence of syntactic segmentation, and SVCs in particular, on event segmentation during perception. In order to isolate this potential influence, it was necessary to also investigate the hypothesized influence of familiarity on event segmentation. This led to the first study to compare event segmentation patterns across cultures and to use the dwell time measure as an independent indicator of event boundaries. Familiarity was indeed shown to influence people's dwell time patterns. However, this influence may have a different effect on short action sequences, which seem to be segmented more finely by those who are familiar with them, in contrast to the longer sequences which are

segmented more coarsely. The use of SVCs, in contrast to coordinate clauses, also appeared to trigger a coarser segmentation in line with their use to refer to single events. This demonstrates the potential for syntactic segmentation patterns to influence the low-level process of ongoing event segmentation during perception.

# 6 Remembering Events of Taking and Putting

An earlier version of the first experiment from this chapter appears as:
Defina, Rebecca & Majid, Asifa. 2012. Conceptual event units of putting and taking in two unrelated languages. In N. Miyake, D. Peebles, & R. Cooper (Eds.), *Proceedings of the 34th Annual Conference of the Cognitive Science Society* (pp. 1470–1475). Austin, TX: Cognitive Science Society.

## 6.1 Introduction

*How was school today? What did you get up to on the weekend? Where did all the ice-cream go?* While we can describe on-going events as we perceive them and talk about things which have not (yet) happened, we often wish to talk about events we have already experienced. In these cases, memory serves as an intermediary between perception and description. Chapter 4 of this thesis showed a close alignment between event units during thinking-for-speaking and syntactic structures, SVCs aligning with single event units. In contrast, Chapter 5 suggested event segmentation during perception is variable, though repeated use of SVCs may encourage a more unified representation than the use of coordinated clauses. What about the representations we form in memory? Is our memory for events already biased towards the representations we use during thinking-for-speaking or is it a more faithful representation of the percept? Previous research suggests possible connections in both directions.

The way we remember events is influenced by the way we segment them during perception. The ability to segment events in a more normative way, i.e. agreeing with the majority of participants, is predictive for how well a person remembers events, even when taking more general cognitive abilities into account (Sargent et al. 2013). An individual's memory for a specific event is also modulated by the way they segment it. Material surrounding perceptual event boundaries is better remembered than material within event segments (Newtson & Engquist 1976; Swallow et al. 2011; Swallow, Zacks & Abrams 2009). Likewise, segmenting events more finely with more event boundaries during perception improves memory for the details of those events overall (Lassiter & Slaw 1991).

Memory for events is also influenced by our understanding of what is happening. So much so that people will fill in remembered events with expected – but unseen – material. For instance, people falsely remember unseen collisions connecting causal

actions and their results in games of badminton (Strickland & Keil 2011). This process is much like the amodal completion of objects, where unseen parts of partially occluded objects are subsequently remembered (Michotte, Thinès & Crabbé 1991).

The way events are described also influences our memory for them. The simple presence of a description can either improve or reduce accuracy depending on whether the description was provided before or after the event (Huff & Schwan 2008). Specific aspects of the description can also alter memory, as in the classic Loftus and Palmer (1974) study. Participants viewed films of car accidents and were asked to estimate the speed of the car when it "smashed/collided/bumped/hit/contacted". The way the event was described in the question influenced people's responses: they reported faster speeds for "smashing" cars than "contacting" cars. Long-term patterns of how people describe events can also influence how they remember them in the absence of overt descriptions. For instance, differences between English and Spanish in the tendency to use agentive descriptions (i.e. "she broke the vase") versus non-agentive descriptions (i.e. "the vase broke") correlate with differences in English and Spanish speakers' memory for agents (Fausey & Boroditsky 2011).

There is, thus, evidence supporting both the bottom-up influence of segmentation during perception and the top-down influence of event concepts and descriptions on memory for events. The present set of studies examines Avatime speakers' memory for taking and putting actions in order to test whether events typically described using SVCs are remembered as single events.

The pattern of describing placement events in Avatime provides an excellent environment for testing whether SVC use influences event memory. Most placements are described with a take-put SVC, as in (1). The first verb in these SVCs is always a verb of taking and the second a verb of placement. The same construction can be used to describe the placement activity alone, or the preceding taking action and the placement action together.

| (1) | *a-kɔ̀* | *kɔ̀ranti=ɛ* | *kpɛ* | *ní* | *kà-sɔ=ya* | *mɛ̀* |
|---|---|---|---|---|---|---|
| | $C_1$S.PFV-take | banana=DEF | put | LOC | $C_6$S-basket=DEF | inside |

'S/he put the banana into the basket.'

There is also a small set of placement events not described with take-put SVCs. These events include putting an article of clothing or jewellery on a body part (in its canonical location), and pouring liquids. They are described using either a put verb in a simple clause (2) or a put verb combined with a manner verb in an SVC (3), respectively. It is strongly dispreferred to describe such actions using a take-put SVC.

(2) *a-kpɛ* *lì-kùto = lè*
$C_1$S.PFV-put $C_3$s-hat = DEF
'S/he put the hat on.'

(3) *e-nyi* *kù-ni = o* *kpɛ* *ní* *kè-zi = a* *mɛ̀*
$C_1$S.PFV-pour $C_5$s-water = DEF put LOC $C_6$s-bowl = DEF inside
'S/he poured the water into the bowl.'

Taking and putting actions provide an ideal situation for testing the relation between SVCs and event memory in Avatime. They constitute a limited domain of separable sequential actions which people perform every day. And, crucially, some of the putting actions are consistently described using SVCs including a taking action, while others are not.

In order to test whether these taking and putting episodes are also remembered as single events, a recognition memory task was designed along the lines of a study by Strickland and Keil (2011). Strickland and Keil showed participants videos of ballistic events, such as a man kicking a football or a woman hitting a shuttlecock with a racket. Some videos showed the trajectory of motion after contact while others did not. Half of the videos showed the moment of contact, while the other half did not. Participants were shown still images from the videos, including the omitted moments of contact. When participants had seen the resulting trajectory of the object, they were very likely to claim they had seen the moment of contact even when they had not. This suggests people mentally fill in parts of events they have not actually seen and will falsely recognise them. If the events described by take-put SVCs in Avatime are remembered as single events encompassing both the taking and the putting action, then seeing one part should lead to false recognition of the other. Thus if an Avatime speaker sees a putting action, which they would describe using a take-put SVC, they should be more likely to falsely recognise a compatible taking action. In contrast, if they see a putting action, which they would not describe with

a take-put SVC, such as putting on clothing or pouring liquid, they should not falsely recognise a corresponding taking action. Note that seeing a taking action is not predicted to trigger false recognition of a putting action as taking actions can always be described using simple single verb clauses and it is the putting action which triggers the use of the SVC.

A control group is needed in order to test whether false recognition rates are related to the use of SVCs rather than inherent differences between events. Crucially, this control group must not combine the taking and putting actions using an SVC or comparable construction. English was initially chosen as the control language (in Experiment 1); however this was changed to Dutch (in Experiment 2) for practical reasons. English and Dutch express putting and taking actions in different ways, for instance Dutch uses positional verbs while English does not (e.g. Alferink & Gullberg 2014). They are, nevertheless, both suitable control languages for this study because neither encodes the taking action as part of the description of a putting action. If a difference in how events are remembered is mediated by the use of SVCs, then it should only be apparent for the Avatime speakers and not the control group. If a difference is apparent for the control group as well, then it is likely driven by more general perceptual or conceptual constraints.

The domain of taking and putting actions is incidentally relevant for a second research question. Several studies have suggested there is a bias towards the goal over the source of motion events. Regier and Zheng (2007) showed people attend more to the goal of putting actions than to the source of taking actions in both language and perception. Papafragou (2010) showed this asymmetry of attention for goals over sources leads to differences in recognition memory: locations and the spatial configurations between them and the figure of motion are remembered more accurately when the location functions as the goal, rather than the source, of a motion event. A corresponding focus on the goal of caused motion events has been reported in an asymmetry in the semantic specificity of putting versus taking verbs cross-linguistically. Languages tend to make more semantic distinctions in the former than in the latter (Narasimhan et al. 2012; Regier 1995). This leads to the prediction that putting events should be remembered more accurately than taking events by speakers of all languages.

Before turning to the experimental data, it is necessary to take a small detour and examine the ways Avatime speakers describe taking and putting actions and the

factors which influence take-put SVC usage in more detail. In particular, it is necessary to understand the motivations for using take-put SVCs to describe placement actions in the first place and the reasons why they are not used in some cases.

## 6.2 Describing putting and taking events in Avatime

The way Avatime speakers describe putting and taking events was initially investigated using an elicitation task. Five native Avatime speakers aged 14-66 (M = 40) were asked to describe what the actor did in 63 short video-clips. Their descriptions were audio recorded and transcribed. The video-clips were those from the Put project developed for investigating descriptions of putting and taking (Bowerman et al. 2004; Narasimhan et al. 2012). Each video-clip was 3-4 seconds long and showed an actor performing either a putting or a taking action. The actions varied along a number of dimensions, such as the nature of the figure, the nature of the ground, the spatial relation between figure and ground, the type of instrument used, and the manner of taking or placing. For a full description of the stimuli see Bowerman et al. (2004) and Narasimhan et al. (2012).

Six placement verbs were used to describe these stimuli, see Table 6.1. There was no generic placement verb, but rather each verb was specified for the type of ground object, the relation between the figure and the ground, and in some cases also the manner of placement or the type of figure object.

Table 6.1: Avatime verbs of putting

| **Avatime verb** | **Semantic characterisation** |
|---|---|
| *kpɛ* | place inside something |
| *trɔ* | place on a flat surface |
| *du* | place on a flat surface, forcefully |
| *plɛ* | place on the ground |
| *sụ* | place an object so that it hangs from something |
| *kume* | place an object which is worn on the head (e.g. glasses or a hat) on the agent's head |

Five take verbs were used, see Table 6.2. In contrast to the putting verbs, one of these was a generic taking verb *kɔ̀* (borrowed from Ewe *kɔ́* 'take, pick up'). Three of

the verbs specify the manner of taking: *feke* is the most general of these and can be used for many situations; *tị* and *halị* are more restricted. The final verb *bu* 'take out' is specified for the spatial relation between the figure and the ground and is the converse of the placement verb *kpɛ* 'put in'.

Table 6.2: Avatime verbs of taking

| **Avatime verb** | **Semantic characterisation** |
|---|---|
| *kɔ̀* | generic taking verb (loan from Ewe) |
| *feke* | take an object, typically where some degree of effort is required (e.g. a bulky or heavy object, or the manner is unusual or difficult) |
| *tị* | pick up an object easily, typically with the finger tips |
| *halị* | take multiple objects at once, typically by gathering them together |
| *bu* | remove object from container |

The most frequent method for describing putting actions in this data was an SVC consisting of a take verb followed by a put verb, as in examples (4) and (5). Twenty-two out of the thirty-four putting actions were consistently described using this construction type.

(4) *a-kɔ̀* *ke-plikpa* *plɛ́* *ke-se=a*
$C_1$s.PFV-take $C_6$s-book put.on.ground:LOC $C_6$s-ground=DEF
'She put a book on the ground.' (Put-Take-007_081031_K)

(5) *lɛ̌* *ɔ-kà=ɛ* *e-feke* *kè-zi=a* *trɔ* *ní*
and $C_1$s-father=DEF $C_1$s.PFV-take $C_6$s-bowl=DEF put.on LOC
*cupi=ye* *abà=ɛ*
cup=DEF on=CM
'And the man put a bowl on the cup.' (Put-Take-031_120922_MM)

Twelve placement actions were not consistently described using a take-put SVC. These fall into three categories:

1) Where an item of clothing is placed on the canonical body part, for instance a hat on a head, see example (6).
2) Where the figure or object placed was the agent's own hand, see example (7).
3) Where a manner verb was used in an SVC with the placement verb in order to modify the manner of placement, see example (8).

(6) *ó-nyime* *e-kume* *lì-kùto = le*
$C_1$s-man $C_1$s.PFV-put.on.head $C_3$s-hat = DEF
'The man put the hat on.' (Put-Take-025_120921_SO)

(7) *o-ne* *a-kpɛ* *a-wla = la* *ní* *ò-se-gù = nɔ*
$C_1$s-mother $C_1$s.PFV-put.in $C_6$s-hand = DEF LOC $C_2$s-tree-stump = DEF
*mɛ* *gì* *lị-bị̀tɛ* *lì-wì = le*
inside CLM $C_3$s.PFV-do $C_3$s-hole = DEF
'The woman put her hand into the tree stump where there was a hole.'
(Put-Take-023_120922_MM)

(8) *a-ŋwya* *ke-plikpa* *plɛ́* *ke-se = a*
$C_1$s.PFV-throw $C_6$s-book put.on.ground:LOC $C_6$s-ground = DEF
'He threw the book onto the ground.' (Put-Take-010_081031_K)

These results converge with data collected outside of the elicitation task. Take-put SVCs are commonly used in Avatime and fall within a more general type of take-SVC where the first verb is a take verb and introduces the direct object, typically the theme, while the second verb describes the action, as in example (9) (see also Section 3.3).

(9) *a-kɔ̀* *lị-ba = lɛ* *kị́ = yɛ*
$C_1$s.PFV-take $C_6$s-hoe = DEF give = $C_1$s.OBJ
'He gave him the hoe.' (Folkstory_110406_QM_05:25)

Such direct object take-SVCs are also commonly found in other serializing languages (Lord 1993). Lord (1993) argues they are used to maintain a pattern of one object per verb. Indeed, this also appears to be the principle motivation in Avatime, where take-SVCs are rarely used with two-place predicates but are very

frequent with three-place predicates. For instance, verbs of placement are inherently three-place predicates with an agent, theme and goal location. When a take-put SVC is used the two objects can be spread out among the two verbs: x takes y puts at z. In most cases where a take-put SVC is not used in Avatime, this one-object-per-verb pattern is maintained in another way. When an item of clothing is placed in its canonical location, the goal location is omitted so only the theme is expressed, as in (6); and when an SVC with a manner verb is used, the theme appears as the direct object of the first verb just as it would in a take-put SVC, as in example (8). It is only cases involving placement of the agent's own hand where a take-put SVC is not used, but the putting verb is nevertheless left with two objects, example (7).

While preserving one-object-per-verb appears to be the primary motivation for using these take-SVC constructions, there are three other factors which influence their use: 1) A take-SVC is more likely to be used with topical objects, such as (10), than with focused objects, such as (11). 2) A take-SVC can also be used to allow zero anaphora of the object, which is otherwise strongly dispreferred in single verb clauses in Avatime, as in (12). This seems to be the main motivation for using a take-SVC with two-place predicates. 3) Take-SVCs can also be used to express intentionality, as in (13).

(10) a. *ɲyaŋwɛ* *wɔ-kɔ̀* *ki-ku=ye* *kɔŋ*
*who* $C_1$S.PFV-take $C_4$S-yam=DEF give
'Who did you give the yam to?'

b. *ma-kɔ̀* *kɨ́* *Komla*
1S.PFV-take give Komla
'I gave (it) to Komla.'

(11) a. *ege* *a-kɨ=yɛ* *na*
*what* $C_1$S.PFV-give=$C_1$S.OBJ QM
'What did he give her?'

b. *a-kɨ=yɛ* *ki-ku=ye*
$C_1$S.PFV-give=$C_1$S.OBJ $C_4$S-yam=DEF
'He gave her the yam.'

(12) A: *wɔ-ŋà* *kɔranti=ɛ*
2s.PFV-eat banana=DEF
'Did you eat the banana?'
B: a. **ee* *ma-ŋà*
yes 1s.PFV-eat
'Yes, I ate (it).'
b. *ee* *ma-kɔ̀* *ŋà*
yes 1s.PFV-took eat
'Yes, I ate (it).'

(13) a. *ma-yrɔ* *katawɛ* *ní* *ke-pe=a* *mɛ̀*
1s.PFV-leave umbrella LOC $C_6$s-yam=DEF inside
'I left the umbrella in the house (perhaps by accident).'
b. *ma-kɔ̀* *katawɛ* *yrɔ* *ní* *ke-pe=a* *mɛ̀*
1s.PFV-take umbrella leave LOC $C_6$s-yam=DEF inside
'I left the umbrella in the house (deliberately).'

In addition to introducing a direct object, Avatime take-SVCs can have two other functions. They can be used to introduce a non-core argument, such as an instrument (14), manner (15) or means (16). They can also be used to indicate that the actor eventually performed the action after long consideration (17)-(18). In contrast to the direct object type, both these functions can only be performed with the generic verb *kɔ̀* 'take'.

(14) *a-kɔ̀* *fork=ye* *feke* *koranti=e* *ní* *ɔ̀-kplɔ=nɔ* *aba*
$C_1$s.PFV-take fork=DEF lift banana=DEF LOC $C_2$s-table=DEF on
'He used the fork (tongs) to lift the banana on the table.'

(15) *a-kɔ̀* *ku-siye=yo* *mè* *sɛ*
$C_1$s.PFV-take $C_5$s-anger=DEF inside leave
'He left in anger.'

(16) *a-kɔ̀* *ku-zi=o* *dzi* *o-honete*
$C_1$s.PFV-take $C_5$s-steal=DEF become $C_1$s-rich.man
'Through theft he became a rich man.'

(17) *a-kɔ̀* *ke-pe = a* *mɛ̀* *si*
$C_1$s.PFV-take $C_6$s-yam = DEF inside paint
'He eventually painted the inside of the house.'

(18) *a-kɔ̀* *o-bi* *pɔ*
$C_1$s.PFV-take $C_1$s-baby give.birth
'She ended up having (keeping) the baby.'

Returning to the take-put SVCs used to describe placement actions, the elicitation task showed these constructions are consistently used to describe most placement events. It also confirmed there is a subset of placement actions for which they are not used. The main motivation for the use of take-put SVCs with placement actions appears to be to preserve the preferred one-object-per-verb strategy. In some related languages, such as Akan, these constructions have grammaticalized as object markers (Lord 1993). This does not appear to be the case in Avatime, however, since different verbs are used depending on the particular taking action, and the object is restricted to things which can be physically taken.

It is also possible to re-examine the co-speech gesture data from Chapter 4 in order to see if it provides any evidence both take and put components are combined into a single event representation. The data from Chapter 4 included nine cases where a take-put SVC was used in combination with a gesture. In most of these (7/9) the gesture overlapped both verbs and referred to the whole transfer from the goal to the source location. In six of these, the gesture simply traced the trajectory of motion and did not specifically refer to taking or putting. In the other, the gesture included clear picking up and releasing motions at the beginning and end of the trajectory. In the remaining two cases, the gesture overlapped either only the placement verb or neither verb and seemed to relate only to the placement action. Despite these two exceptions, the main pattern suggests both the taking and putting actions are active and form a single event unit during thinking-for-speaking.

The following experiments test whether these taking and putting actions are also remembered as single events. The first experiment tested participants' memory in a non-linguistic setting before any verbal descriptions of the actions to see whether people are generally more likely to remember taking and putting actions described with SVCs as single events. The second experiment examined the influence of

describing the actions before performing the memory task to see whether they remember them as single events after explicitly describing them with SVCs. Finally, the third experiment tested the influence of inhibiting linguistic encoding using verbal interference to see whether any tendency to remember the actions as single events was modulated by (possibly covert) language use.

## 6.3 Experiment 1

Avatime and English participants viewed either the taking or putting action from a take-put episode. Half of these episodes would be described in Avatime using take-put SVCs, while the other half would not. They later saw these actions again as well as the corresponding putting or taking actions they had not previously seen. If Avatime speakers remember the SVC-type take-put episodes as single events, they should be more likely to falsely recognise new taking actions if they correspond to a previously seen SVC-type putting action. If such a difference is linked to the use of take-put SVCs rather than differences in the actions themselves, it should be observed among Avatime but not English speakers.

### 6.3.1 Method

#### 6.3.1.1 Participants

Thirty-four native speakers of Avatime, aged 11-16 (mean 14.1 years), were recruited at Vane Junior High School, Ghana. Four Avatime speakers were excluded due to technical difficulties or for uniformly answering either yes or no to all items. Thirty-three native speakers of English, aged 11-17 (mean 14.2 years), were recruited in the Blue Mountains and Sydney, NSW, Australia.

All Avatime speakers were also fluent in Ewe and English and 11 additionally spoke Twi. Ewe and Twi are also serializing languages and make use of similar take-put constructions to describe placement actions (Ameka 2006a; Lord 1993). In both Ewe and Twi, however, the ‘take’ verb has been semantically bleached. In Ewe, it is no longer possible to vary the ‘take’ verb according to the manner of taking and in Twi the ‘take’ verb has fully grammaticalized as a patient marker (Lord 1993). English does not make use of any such take-put constructions. Bilingualism in any of these languages could thus decrease the chances of finding the predicted effect. In order to minimize any impact from these languages, all efforts were made to maintain Avatime as the dominant language throughout the experiment.

There was much less bilingualism among the English participants and none of the other languages spoken by the participants are expected to interfere with the study since, like English, they do not make use of take-put constructions. One English speaker was also fluent in German. Two additionally spoke Spanish, one fluently and the other moderately. Of the remaining English speakers, 9 were completely monolingual and 21 had very limited knowledge of another language (French, German, Italian, Japanese, Korean or Latin).

#### 6.3.1.2 Materials

80 paired putting and taking actions were filmed in a single location. They were performed by two actors, one male and one female. Each video-clip lasted 3-4 seconds. Each take video-clip showed one actor removing an object from a location. The corresponding put video-clip showed the same actor placing the same object in a different location. For instance, in one pair a man takes a banana from the shelf and places it on a plate (Figure 6.1(a)), in another a woman takes a necklace from a bag and places it around her neck (Figure 6.1(b)). Within these paired taking and putting episodes, the camera angle and position of the actor in the room were kept constant.

The objects and locations were selected to be familiar for both Avatime and English speakers. The source of the taking action was always different from the goal of the putting action. Across episodes, the object, locations, position of the actors, and camera angle varied.

Half of the 40 take-put episodes had general placement actions of the type described using take-put SVCs in Avatime. The other half showed placements such as the donning of clothing and pouring of liquids, which are not described using take-put SVCs in Avatime. Descriptions of the items by Avatime participants at the end of the experiment confirmed this distinction: 96.2% (SD = 1.8) of the placement actions in the SVC category were described using take-put SVCs, in contrast only 6.5% (SD = 1.7) of the Non-SVC placement actions were described using a take-put SVC. This difference was significant $t(38) = 48.79$, $p < .001$. For ease of reference, both putting and taking actions in an episode will be referred to as either SVC or Non-SVC according to the type of putting action.

a.

b.

Figure 6.1: Still images from two pairs of stimuli. The images on the left come from the taking actions, the images on the right come from the putting actions. Thus (a) shows the man taking a banana from a shelf (left) and putting it on a plate, which Avatime speakers would describe using an SVC (right). In (b) the woman takes a necklace from a bag (left) and places it around her neck, which Avatime speakers would not describe using an SVC (right).

#### 6.3.1.3 Design

The 40 take-put episodes resulted in 80 individual items each consisting of a single take or put action. Participants saw either the taking or the putting action from each take-put episode in a learning phase, before seeing all items – those they had already seen and their putting or taking counterparts – in a testing phase. Pilot testing with Avatime speakers showed remembering all 40 learning items in one go was too difficult, so the experiment was divided into two blocks. In each learning block, there were 5 SVC put actions, 5 Non-SVC put actions, 5 SVC take actions, and 5 Non-SVC take actions. Block order was counterbalanced across participants. Within each block, items appeared in one of four random orders.

#### 6.3.1.4 Procedure

Participants were asked to watch a series of video-clips and to remember them as best they could. They were told they would later be shown more videos, some the same as those they had seen and some different. Their task was to say which were the same and which different. Instructions and responses were given verbally in the participant's native language. Responses were audio recorded.

Following the explanation there was a training task to confirm participants understood the task. They initially saw three video-clips and were then tested on a set of six: three the same as they had just seen and three new items. Two of the new items showed the actor performing a very different action and one showed only a small variation in the action performed.

Participants watched video-clips one at a time during the learning phase. The video-clips were separated by a black screen lasting 1 second. There was then a 5 minute distraction task unrelated to the experiment (free listing of items in semantic domains, such as animals or body parts). Participants were then tested for their memory of the 20 video-clips they had just seen, plus their 20 unseen counterparts. So, if a participant saw a woman put on a necklace in the learning phase, they now, in the testing phase, also saw the woman taking the necklace out of the bag. Participants indicated whether each video-clip was the same or different to one they had seen previously. The same procedure was then repeated with the second block.

After the memory experiment, participants viewed all video-clips again and were asked to describe 'what the person did'.

Participants were tested individually and the same procedure was used for English and Avatime participants. The whole experiment lasted approximately 45 minutes.

### 6.3.2 Results and discussion

Responses to previously seen and new items were analysed separately using logistic mixed effects with the lme4 package (Bates, Maechler, et al. 2015) in R 3.2.1 (R Core Team 2015). This method of analysis allows the influence of experimental conditions on the dependent variable, in this case whether or not the response was correct, to be tested while also accounting for variation between participants and stimuli items simultaneously. Following Barr et al.'s (2013) recommendations on avoiding anticonservative biases in mixed effects models, the maximal model was

used as the starting point for analysis. This included all fixed effects – Language (Avatime or English), Construction-type (SVC or Non-SVC), and Action-type (take or put) – as well as all interactions between them, random intercepts for participants and stimuli items, and random slopes for each fixed effect term. This initial model was then reduced as follows. Random slope terms were removed according to the algorithm proposed by Bates, Kliegl, Vasishth, and Baayen (2015) only when their removal did not significantly reduce the goodness-of-fit of the model and was necessary for achieving convergence. Fixed effect terms were removed in order of least significance as long as their removal did not result in a near significant ($p < .100$) change in the goodness-of-fit or impact the convergence of the model.

Responses to the previously seen items were analysed first to determine whether participants were able to correctly indicate that they had seen the video-clip when they had in fact previously seen it. There is no specific hypothesis for these previously seen items concerning the SVC distinction, though putting actions are predicted to be recognised more accurately than taking actions following the source vs. goal asymmetry. The model reduced to include a two-way interaction between Language and Construction-type and a main effect of Action-type, with a random slope for Language only. Both groups showed sufficiently high levels of recognition, though English speakers recognised the previously seen items more accurately (M = 83.56%) than Avatime speakers (M = 80.67%) $\beta = 0.66$, SE = 0.27, $z = 2.46$, $p = .014$. There was a significant interaction between Language and Construction-type: Surprisingly, English speakers recognised Non-SVC actions more accurately (M = 87.27%) than SVC actions (M = 79.85%), while Avatime speakers showed little difference (M = 81.50% vs. M = 79.83%) $\beta = -0.50$, SE = 0.25, $z = -1.98$, $p = .048$. Construction-type was not significant as a main effect $\beta = -0.12$, SE = 0.17, $z = -0.72$, $p = .472$. As predicted by the source vs. goal asymmetry, putting actions were recognised more accurately (M = 85.08%) than taking actions (M = 78.73%) by speakers of both languages $\beta = -0.44$, SE = 0.16, $z = -2.75$, $p = .006$.

The primary hypothesis concerned the new items. Here, the model was reduced by removing the random slopes for Construction- and Action-type. No fixed effects could be removed without reducing the convergence of the model. The prediction was that Avatime speakers would be more likely to falsely recognise new taking actions when they corresponded to previously seen SVC, rather than Non-SVC, putting actions. This prediction was not borne out, since the interaction between Language, Construction- and Action-type was not significant $\beta = 0.18$, SE = 0.48,

$z = 0.37$, $p = .711$. The only significant effect among the new items was Language: Avatime speakers had more false recognitions (M = 28.33%) than English speakers (M = 15.77%) $\beta = 0.92$, SE = 0.32, $z = 2.89$, $p = .004$.

There was thus no evidence of the predicted tendency for Avatime speakers to falsely recognise new Non-SVC take actions. In fact, the only difference in response to SVC and Non-SVC actions was shown by the English speakers who recognised previously seen Non-SVC actions more accurately than SVC ones. This suggests English speakers are responding to general differences between the actions which Avatime people describe using take-put SVCs and the ones they do not. It is, however, striking that the English speakers show this difference while the Avatime speakers do not.

## 6.4 Experiment 2 – Linguistic Description

This experiment tested whether overtly describing actions would influence people's recognition memory for them. Previous work (e.g. Loftus & Palmer 1974) has demonstrated the influence of explicit verbal descriptions on event memory, so Avatime speakers may be more likely to remember actions as single events after explicitly describing them with an SVC. Avatime and Dutch speakers performed the same recognition memory task from Experiment 1, with an added comparison between participants who described the actions in the learning phase and those who did not. The hypothesis that Avatime people remember the actions described by take-put SVCs as single events predicts Avatime speakers will falsely recognise new SVC take actions, at least after they have overtly described the corresponding putting action with a take-put SVC. Alternatively, if there is a more universal difference in how people remember the SVC and Non-SVC actions, then Dutch speakers should also respond differently to the SVC and Non-SVC actions.

### 6.4.1 Method

#### 6.4.1.1 Participants

Two new sets of participants were recruited, none of whom had participated in the first experiment. Fifty-three native speakers of Avatime, aged 12-19 (mean 14.0 years), were recruited at Amedzofe Junior High School, Ghana. Eight participants did not proceed with the experiment because they did not complete the training phase correctly. Eleven participants were tested, but later excluded due to technical

difficulties or for uniformly answering either yes or no for all items. This left thirty-four participants, aged 12-19 (mean 14.2 years). As in the previous experiments, there was a great deal of bilingualism among the Avatime participants. They were all fluent in Ewe and, with one exception, also spoke English. Sixteen additionally spoke Twi.

Sixty-nine native speakers of Dutch, aged 12-17 (mean 13.7 years), were recruited from Pax Christi college in Druten, Olympus college in Arnhem, and via advertisements in Nijmegen. Five participants were tested but excluded due to technical difficulties. Just as with the Avatime participants, Dutch participants were also largely bi- or multi-lingual. They reported knowledge of the following languages (number of participants given in brackets): English (54), French (26), German (24), Spanish (4), Indonesian (1), Latin (1), Macedonian (1), Mandarin (1), and an unidentified language of the Maluku province of Indonesia (1). While there is a much greater degree of multilingualism among the Dutch participants here than among the English participants in Experiment 1, only the last two languages present possible confounds: Mandarin also makes use of take-put constructions similar to Avatime and the behaviour of the language from Maluku is unknown. These two participants were thus excluded from the analysis. This left thirty-one Dutch speakers who described the actions, and thirty-one who did not.

#### 6.4.1.2 Materials

The 80 video-clips from Experiment 1 (Section 6.3.1.2).

#### 6.4.1.3 Design

Dutch participants were randomly assigned to either the Describe or No-Describe condition. All new Avatime participants were placed in the Describe condition and the Avatime responses from Experiment 1 were used for the No-Describe condition. As in Experiment 1, participants saw either the take or put action from each episode in a learning phase (the seen set) and were later tested on their recognition of all items (previously seen and new). Testing was again divided into two blocks. In each learning block, there were 5 SVC put items, 5 Non-SVC put items, 5 SVC take items, and 5 Non-SVC take items. Each testing block consisted of the 20 learning items plus their take or put counterparts. Block order was counterbalanced across participants. Within each block, items were presented in a random order.

#### 6.4.1.4 Procedure

Participants were asked to watch a series of video-clips and remember them as best they could. They were told they would later be shown more videos and their task would be to indicate which were the same as those seen previously, and which were new. Those in the Describe condition were asked to describe what the person did after each item in the learning phase. Those in the No-Describe condition described the video-clips only after completing the memory tests. Instructions were given verbally in the participant's native language. Description responses were given verbally and recorded. Recognition responses were given via a button box: participants pressed a green button on the right to indicate the item was the same as one in the learning phase and a red button on the left to indicate it was different.

As in Experiment 1, there was a short training phase following the instructions. This training phase was extended to give participants practice using the button box. There were 5 video-clips in the learning phase and 10 in the testing phase. Some of the new video-clips were very different from the learning items with different actors and locations as well as different actions. Participants were considered to have successfully completed the training if they answered at least 7 out of the 10 items correctly.

Following each item in the learning phase, participants in the No-Describe condition saw a black screen for 1 second before the next item began automatically. Participants in the Describe condition saw a blank, blue screen while they described the preceding action. The experimenter manually triggered the next video-clip when they had finished.

After the learning phase, there was a five minute distraction task unrelated to the experiment. Participants in the Describe condition were asked to repeat numbers and rhythms (so as to test the appropriateness of the interference materials for Experiment 3). Participants in the No-Describe condition performed the same semantic domain listing task as in Experiment 1.

Participants were then tested for their memory of the first block of 20 video-clips they had just seen, plus their previously unseen counterparts. After completing the first block, the procedure was repeated for the second block.

Participants were tested individually and the same procedure was used for Dutch and Avatime participants. The whole experiment lasted approximately 30 minutes in the Describe condition and 45 minutes in the No-Describe condition. This

difference in duration arises from the repetition of all items so participants in the No-Describe condition can describe the videos after completing the recognition task.

### 6.4.2 Results and discussion

The responses were analysed using logistic mixed effects as in Experiment 1 with responses to seen and new items analysed separately. The dependent variable was whether or not the response was correct. The initial model included all experimental conditions (Language, Description condition, Action-type, and Construction-type) with all possible interactions, random effects of participant and stimuli item and random slopes for each experimental condition. This model was reduced following the same algorithm as Experiment 1 to arrive at the final models.

The final model for the previously seen items included two-way interactions between Language and Description condition and between Description condition and Action-type, as well as a main effect of Construction-type. All random slope terms were removed to reach convergence. Both Avatime and Dutch speakers reliably recognised the video-clips they had previously seen (M = 86.76% vs. M = 90.52%) and there was no difference between the two language groups $\beta = 0.07$, SE = 0.22, $z = 0.30$, $p = .763$. People who described the actions during the learning phase recognised them more accurately (M = 92.38%) than those who did not (M = 84.59%) $\beta = -1.30$, SE = 0.24, $z = -5.45$, $p < .001$. Putting actions were also recognised more accurately (M = 91.47%) than taking actions (M = 85.75%) $\beta = -0.85$, SE = 0.20, $z = -4.49$, $p < .001$. This corroborates Experiment 1 in confirming the source vs. goal based prediction that putting actions would be recognised more accurately than taking actions. There were no other significant effects. In particular, there were no significant differences according to Construction-type $\beta = -0.23$, SE = 0.14, $z = -1.66$, $p = .097$.

For the new items, the final model included a two-way interaction between Description condition and Action-type as well as main effects of Language and Construction-type. All random slopes, except for Language, were removed. The hypothesis predicted Avatime people would falsely recognise taking actions when they corresponded to previously seen SVC putting actions, at least when they had described the actions. However, both the four-way interaction ($\beta = -0.72$, SE = 0.68, $z = -1.06$, $p = .290$) and the three-way interaction between Language, Action- and Construction-type were not significant ($\beta = -0.27$, SE = 0.41, $z = -0.66$, $p = .507$) and removed as part of the model reduction process. Avatime speakers had more false

recognitions (M = 26.64%) than Dutch speakers (M = 9.80%) β = 1.43, SE = 0.17, $z$ = 8.40, $p$ < .001. People who described the actions during the learning phase were less likely to make false recognitions β = -0.61, SE = 0.17, $z$ = -3.48, $p$ < .001. However, this difference was only apparent in response to putting actions (M = 15.23% vs. M = 21.64%), not taking actions (M = 18.39 vs. M = 18.36%) β = 0.52, SE = 0.16, $z$ = 3.25, $p$ = .001. Finally, both Avatime and Dutch speakers had more false recognitions for SVC- (M = 20.44%) than Non-SVC-type actions (M = 16.27%) β = -0.33, SE = 0.16, $z$ = -2.02, $p$ = .043.

Although these results differ from those of Experiment 1, they are in line with the hypothesis that there are general differences in how people remember these SVC and Non-SVC actions. In the first experiment, English speakers recognised previously seen Non-SVC actions more accurately than SVC actions. Here the difference is among the new items, where both Avatime and Dutch speakers had less false recognitions for Non-SVC actions than SVC actions. In both experiments, people were more accurate with the Non-SVC actions. In neither experiment was there evidence of a difference mediated by the Avatime speakers' use of take-put SVCs. Rather it seems that the marked manners and actions directed towards the agent's own body which lead Avatime people not to use take-put SVCs also make these actions more universally memorable.

## 6.5 Experiment 3 - Interference

This experiment was designed to test whether covertly describing the actions during the learning phase could influence how Avatime people recognise the SVC and Non-SVC putting and taking actions. A new set of Avatime participants were tested after performing a verbal interference task during the learning phase. This dual task inhibits covert linguistic encoding and so should reduce any linguistically mediated differences in the recognition accuracy of SVC versus Non-SVC or putting versus taking actions. In order to ensure any changes were due to inhibition of linguistic encoding rather than simply the demands of performing a dual task, another group of Avatime participants were tested after performing a comparable non-verbal interference task. There was no need to test how English or Dutch speakers performed under verbal interference, since they do not make use of the relevant linguistic construction and the comparison with Avatime speakers performing a non-verbal interference task provides a sufficient control group.

### 6.5.1 Method

#### 6.5.1.1 Participants

Ninety-six native speakers of Avatime, aged 10-20 (mean 14.7 years), were recruited at Junior High Schools in Amedzofe, Biakpa and Gbadzeme. None of these participants had taken part in either Experiment 1 or 2. Twenty-three participants did not proceed with the experiment because they did not complete the training phase correctly. Thirteen participants were excluded for answering either yes or no to all items, not performing the interference task correctly, not attending to the experiment, or leaving before completing the experiment. This left sixty participants (mean 14.5 years) whose responses were included in the study. All participants, except one, additionally spoke English and Ewe. Fifteen additionally spoke Twi, two Ga, one French, and another Likpe (a Ghana-Togo Mountain language spoken a little to the north of Avatime).

#### 6.5.1.2 Materials

The video-clips were the same as those in Experiment 1. There were two sets of interference stimuli: verbal and non-verbal. Both sets consisted of audio recordings lasting the full duration of the video-clips (3 seconds).

The verbal interference stimuli consisted of 40 recordings of a male Avatime native speaker reading a random sequence of three of the Avatime numerals one through to five. From experience in the community, I found the Avatime numerals above five were not well known by younger speakers, who more often used English for higher numbers. The Avatime numerals one through five are all disyllabic: *ole*, *ɔba*, *ɔta*, *one*, *otsu*, so sequences of three numerals could be said slowly and clearly within the 3 second duration of the video-clip. The pronunciation of the numerals varies between the different Avatime dialects so different recordings were used to suit the local pronunciation for each testing site.

The non-verbal interference stimuli consisted of 20 computer-generated rhythms. Pilot testing showed generating 40 distinct rhythms required them to be too complex to be accurately repeated by participants. Thus, two sets of 20 distinct rhythms were generated using different sound effects (a clap and a cow-bell). One set was played with the items in Block A and the other with Block B so participants heard only distinct rhythms within each block. Rhythms provide a good non-verbal comparison since they also involve remembering a sequential auditorily presented pattern, but they do not involve linguistic items or vocalisation. Similar rhythmic

interference tasks were also used for non-verbal interference conditions by Trueswell and Papafragou (2010) and Newton and de Villiers (2007).

Avatime speakers' ability to repeat both the verbal and non-verbal interference materials was tested separately, during the distraction task in Experiment 2. They were very accurate with both the verbal (M = 93.3%) and the non-verbal stimuli (M = 95.8%). There was no significant difference between the two $t(52) = 1.44$ $p = .157$. This suggests it would be possible for Avatime speakers to perform either task while still attending to the video-clips for the recognition task.

#### 6.5.1.3 Design

Participants were randomly assigned to either the Verbal or Non-Verbal interference condition. The responses from the Avatime speakers in Experiment 1 were also included in the analysis in order to contrast the dual task interference conditions with a no interference condition. As in Experiment 1, participants saw either the take or put action from each episode in the learning phase and were later tested on their recognition of all items. Testing was again divided into two blocks with 5 SVC put, 5 Non-SVC put, 5 SVC take, and 5 Non-SVC take items in each block. Each testing block consisted of 20 learning items plus their put or take counterparts. Blocks were counterbalanced across participants and items were presented in a random order within each block.

#### 6.5.1.4 Procedure

As in Experiment 1, participants were asked to remember video-clips so they could say whether the test items were the same or different. In addition, they were asked to listen to the interference items (numbers or rhythms) and repeat them after each item (verbally or by tapping on the table). Instructions were given verbally in Avatime. Recognition responses were given via a button box, participants pressed a green button on the right to indicate an item was the same as one in the learning phase and a red button on the left to indicate it was different.

The extended training phase from Experiment 2 was also used here to give participants practice using the button box to respond. This training phase was further extended by allowing participants to repeat the learning phase until they felt confident performing the dual task.

With the exception of the requirements of the dual task, the rest of the experiment continued as in Experiment 1. Participants again listed items from semantic domains for 5 minutes between the learning and testing phases as a distraction task. The whole experiment lasted approximately 45 minutes.

### 6.5.2 Results and discussion

Responses were analysed using logistic mixed effects as in previous experiments. The dependent variable was again whether or not the response was correct. The initial model included all experimental conditions – Interference condition (verbal, non-verbal, or none), Action-type (put or take), and Construction-type (SVC or Non-SVC) – with all possible interactions, random effects of participant and stimuli item and random slopes for each experimental condition.

For the previously seen items, the model was reduced to include a two-way interaction between Action- and Construction-type and a main effect of Interference condition. The random slope for Interference condition was removed to achieve convergence. Participants in all Interference conditions were able to recognise the video-clips they had previously seen. Although, those who performed the verbal (M = 72.67%) and non-verbal interference tasks (M = 70.58%) were less accurate than those who performed no dual-task (M = 80.67%) $\beta = -0.57$, SE = 0.20, $z = -2.81$, $p = .005$ and $\beta = -0.61$, SE = 0.20, $z = -3.01$, $p = .003$, respectively. As in the previous experiments, putting actions were recognised more accurately (M = 79.11%) than taking actions (M = 70.17%) $\beta = -0.87$, SE = 0.21, $z = -4.15$, $p < .001$. Also just as the English speakers in Experiment 1, the Avatime speakers here showed a tendency to recognise Non-SVC actions more accurately (M = 75.83%) than SVC ones (M = 73.44%) $\beta = -0.56$, SE = 0.21, $z = -2.70$, $p = .007$. However, this difference between Non-SVC and SVC actions was only apparent among the putting actions (M = 82.11% vs. M = 76.11%), not among the taking actions (M = 69.55% vs. M = 70.78%) $\beta = 0.56$, SE = 0.28, $z = 2.04$, $p = .042$. This restriction to putting actions is consistent with the fact that the distinction between SVC and Non-SVC take-put action pairs is motivated by the properties of the putting action. There were no other significant effects.

For the new items, the model reduced to include two-way interactions between Interference condition and Construction-type and between Interference condition and Action-type, the random slopes for Interference and Construction-type were removed leaving only the slope for Action-type. As in Experiment 2, people were

more likely to falsely recognise SVC actions (M = 32.94%) than Non-SVC actions (M = 29.56%) β = -0.41, SE = 0.19, $z$ = -2.21, $p$ = .027. Unexpectedly, this difference was only apparent for those who performed either the verbal interference task (M = 30.50% vs. M = 26.33%) or no interference task (M = 31.67% vs. M = 25%). It was not apparent for people who performed the non-verbal, rhythmic interference task (M = 36.67% vs. M = 37.33%) and this interaction between Interference condition and Construction-type was significant β = 0.43, SE = 0.19, $z$ = 2.29, $p$ = .022. Those who performed the non-verbal rhythmic task also diverged from the others in their responses according to Action-type β = -0.61, SE = 0.22, $z$ = -2.82, $p$ = .005. In both the no interference and verbal interference conditions people had more false recognitions for putting rather than taking actions (M = 30.83% vs. M = 25.83% and M = 30.17% vs. M = 26.67%, respectively). However, in the non-verbal interference condition people had more false recognitions for taking (M = 40.00%) rather than putting actions (M = 34.00%).

Contrary to the initial prediction, the pattern of responses from people completing the verbal interference task did not diverge from those who performed no interference task. This indicates covert linguistic descriptions were not motivating the Avatime people's false recognitions of SVC-type actions. This is in line with the results of the first two experiments which also suggested the differences between how SVC and Non-SVC actions are recognised were due to more general, non-linguistic, cognitive factors.

Interestingly, the responses of those who performed the non-verbal interference condition did diverge from the others. One possible explanation is that this dual-task encouraged people to make more use of linguistic encoding as a memory aiding strategy. This possibility was suggested by Trueswell and Papafragou (2010) to explain their finding that people paid more attention to elements which were encoded in their language when the task was made more difficult by adding either a rhythmic interference or delayed counting task. This explanation would fit the present finding that people who performed the non-verbal interference task were more accurate in rejecting new putting rather than taking actions since putting actions are encoded more specifically in Avatime. It would not however, explain why these participants showed similar rates of false recognitions for SVC- and Non-SVC-type actions. An alternative explanation could be that the rhythmic task triggered some unexpected disturbance in how these actions are remembered.

## 6.6 General discussion

This set of studies set out to investigate whether Avatime speakers remember the actions they describe with take-put SVCs as single events. The studies were based on the rationale (supported by Strickland and Keil, 2011) that if Avatime speakers remember SVC-type take-put episodes as single events, seeing an SVC-type putting action should trigger false-recognitions of the taking action. The results provided no evidence of such SVC induced false recognition. Rather, the results suggest a general tendency for people to remember the actions which Avatime people do not describe with take-put SVCs more accurately than the ones which they do.

In all three experiments people responded more accurately to Non-SVC actions than SVC ones. However, this difference appeared only in some conditions, with the specific distribution varying across the experiments. In the first experiment, English speakers responded more accurately to previously seen Non-SVC actions. In the second experiment, both Avatime and Dutch speakers had less false recognitions, i.e. more accurate rejections for new Non-SVC actions. In the third experiment, Avatime speakers across the different interference conditions responded more accurately to previously seen Non-SVC putting actions than to SVC putting actions. Those who performed either the verbal interference task or no interference task also responded more accurately to new Non-SVC items than SVC ones. The overall pattern confirms there is a difference in how people remember these kinds of actions and that this difference is not limited to the new taking actions nor to the Avatime speakers as was predicted according to an account based on the use of take-put SVCs. Speakers of Dutch and English also responded more accurately to Non-SVC actions than SVC ones, even though they do not use take-put SVCs or any other linguistic means to mark a categorical distinction between these two groups of actions. This along with the lack of influence from explicit verbal encoding or interference suggests general non-linguistic properties of the actions themselves, rather than Avatime speakers' use of SVCs, influence how people remember them. The tendency to recognise Non-SVC items more accurately than SVC items thus appears to be a characteristic of thought which produces a 'thought on language' effect (e.g. Gerrig & Banaji 1994; Regier, Kay & Khetarpal 2007) in the case of how Avatime people use take-put SVCs.

The most likely explanation is that the very same action characteristics drive both the linguistic distinction in Avatime and the accuracy differences among speakers of Avatime and other languages. As discussed in Section 6.2, putting actions are by default described using take-put SVCs in Avatime. There are three

kinds of actions which are exempt from this general requirement: 1) those where the manner of placement warrants further description using another manner verb (e.g. pouring water into a glass); 2) those where the object is an item of clothing or jewellery and is placed on the agent in the typical location (e.g. placing a hat on one's head); or 3) those where the object placed is the agent's own hand (e.g. putting one's hand into a bag). The present set of experiments included examples from the first two groups as stimulus items in the Non-SVC putting action condition. These characteristics – marked manners and actions on the agent's own self – could also lead people to pay more attention to these actions. This would explain why people responded more accurately to both the previously seen and new Non-SVC putting actions than the SVC-type ones. Better memory for previously seen Non-SVC-type putting actions could also help people to accurately reject the corresponding new taking actions. Such effects have been reported in other recognition memory studies where people are better able to discriminate deviations from the events for which they create more detailed (e.g. Papafragou 2010) or task suitable (e.g. Huff & Schwan 2008) memory representations. This would explain why people tended to more accurately reject both putting and taking Non-SVC-type actions.

The general tendency to recognise Non-SVC items more accurately than SVC items makes it impossible to determine – on the basis of the current experiments – whether Avatime speakers remember the SVC taking and putting actions as single events. It is possible that remembering these events as single units serves as an additional factor in leading Avatime speakers to incorrectly recognise new SVC items. Indeed, it is very likely that Avatime speakers would have such a false-recognition effect, at least following explicit verbal encoding. Priming with SVCs was associated with a more unified event representation during perception in Chapter 5 and several studies have shown that event units in memory are strongly influenced by the boundaries formed during perception (e.g. Swallow, Zacks & Abrams 2009). Studies on the influence of verbal encoding on event memory (e.g. Huff & Schwan 2008; Loftus & Palmer 1974) also suggest that the use of take-put SVCs would make Avatime speakers more likely to falsely recognise a taking action corresponding to an SVC-type putting action simply because they would have produced a description of the taking action as part of their description of the putting action. The fact that the Avatime speakers tended to have a larger difference in

accuracy between the new Non-SVC and SVC items than between the previously seen ones, is in line with such an explanation. However, one would also expect such an effect to interact with explicit verbal description and interference, and there was no evidence of this. The possibility thus remains open and would need to be tested in another experiment which controlled for the unpredicted general asymmetry found between Non-SVC and SVC putting actions. If it turns out that these take-put SVCs do not in fact lead Avatime speakers to falsely-recognise the corresponding taking actions then perhaps the take verbs in these SVCs are more grammaticalized and semantically bleached than they appear. One way to test this would be to look at co-speech gesture patterns occurring with these take-put SVCs. The few examples found among the data used in Chapter 4 – discussed in Section 6.2 – suggested the semantics of the taking verb remain active in the SVC.

A secondary research question concerned whether putting events are remembered more accurately than taking events. Previous research has shown people are better able to discriminate differences in objects and spatial orientations when they function as the goal rather than the source of a motion event (Papafragou 2010; Regier & Zheng 2007). It was thus predicted people would respond more accurately to the goal-oriented putting actions than to the source-oriented taking actions. Such an asymmetry was found, but only among previously seen items. While it is very possible this asymmetry applies only to previously seen actions, it is surprising given that previous studies also report a source vs. goal asymmetry among new actions (Papafragou 2010; Regier & Zheng 2007). Another possible explanation for the restriction found here is that the pairing of taking and putting actions in these experiments may have lead people to be just as accurate in rejecting the taking actions corresponding to putting actions they had previously seen as they were in rejecting new putting actions.

## 6.7 Conclusion

The goal of this chapter was to investigate the way events are discretized in memory. Specifically, whether Avatime speakers remember the actions they describe using take-put SVCs as single events. Instead of clear evidence to either support or contradict this, the studies in this chapter uncovered a general bias towards remembering actions which Avatime speakers do not describe using take-put SVCs more accurately than those which they do. This asymmetry was shown by Dutch and English as well as Avatime speakers, even though they mark no distinction

between these two groups of actions. These studies have thus uncovered a general distinction in how people think about these events which appears to drive the linguistic distinction in Avatime: an influence of thought on language.

# 7 Conclusions

## 7.1 Summary

This thesis set out to investigate the relations between the ways events are encoded linguistically and conceptually. In particular, it focused on serial verb constructions (SVCs) and the claim that their multiple verbs work together to refer to a single conceptual event. The approach featured a detailed study within one language – Avatime – and the use of multiple methods to investigate different aspects of SVCs and their relationships with conceptual event units during perception, memory, and thinking-for-speaking.

The focus on one particular construction type within one language allowed the investigation to be well-grounded in linguistic analysis. Towards this end, Chapter 2 provided a general sketch grammar of Avatime to inform and situate the discussion of SVCs within the wider language system. Chapter 3 focused on the SVCs in more detail. I began by describing their properties and how they can be distinguished from other Avatime construction types. I then discussed their different semantic functions and what kinds of events they describe. This description uncovered three distinct subtypes of SVC in Avatime: nuclear, core, and sequential. These subtypes are distinguished by their morphosyntactic properties, but also strongly linked to different semantic functions.

This linguistic description was not only essential grounding for the rest of the thesis, but also of wider relevance as a substantial contribution to the description of Avatime. The description shows Avatime SVCs have many properties in common with those of other languages in the region, for instance allowing independent aspect marking on verbs within SVCs, and not allowing switch-subject type constructions. However, Avatime SVCs also show some peculiarities, most notably the distinction between nuclear and core subtypes which is commonly seen in South-East Asia and Oceania, but not West Africa. Avatime SVCs are also notable for making use of a very rare form of agreement marking where a reduced form of subject agreement is used with non-initial verbs in SVCs.

The experimental part of the thesis took a multi-method approach to investigating the relationship between Avatime SVCs and conceptual event segmentation at three different stages: thinking-for-speaking (Chapter 4), perception (Chapter 5), and memory (Chapter 6).

The study in Chapter 4 used co-speech gestures as an indication of event segmentation during thinking-for-speaking. The alignment of event-related co-

speech gestures with SVCs and other multi-verbal constructions was compared across four monologues. SVCs often occurred with single gestures overlapping the entire construction, but never with multiple distinct gestures. In contrast, other multi-verbal constructions often occurred with multiple gestures and only rarely with single gestures overlapping more than one verb. This suggests SVCs do indeed refer to single conceptual events, at least during thinking-for-speaking.

Chapter 5 explored event segmentation in its earliest stages, during perception. I first addressed the question of whether events typically described using SVCs are inherently more likely to be more holistically segmented during perception due to their familiarity. Sequential actions are only described using SVCs if they constitute culturally recognisable units. Since familiar events have been predicted to be segmented more holistically than unfamiliar ones (e.g. Hard, Tversky & Lang 2006; Heider 1958; Zacks et al. 2007; Zacks, Tversky & Iyer 2001), greater familiarity may be a possible motivation for more holistic segmentation of the events typically described with SVCs. Two experiments investigated this possibility by comparing the segmentation of culturally familiar and unfamiliar events by Avatime and Dutch people. A familiarity difference was found only in the experiment where participants described each event immediately after viewing. This familiarity difference was unexpectedly split according to the length of the stimulus. Longer events showed the predicted difference where the people who were familiar with the event segmented it more coarsely than those for whom it was unfamiliar. However, shorter events showed the opposite pattern: they were segmented more finely by those familiar with them. There are a few possible explanations for this pattern of results, each suggesting possible modifications to our understanding of how we segment events during perception. However, the take-home message for the present investigation is that while familiarity does seem to influence the way people segment events it is not likely to lead to an inherent tendency to holistically segment the events typically described with SVCs.

The second main question in this chapter was whether the use of SVCs would influence the way people segment events during perception. Avatime-speaking participants were primed either with sequential type SVCs or their coordinate clause paraphrases before segmenting the Avatime- and Dutch-familiar events from the previous studies in this chapter. Syntactic priming did indeed influence segmentation: Participants primed with SVCs had fewer fine-level event boundaries

than those primed with coordinate clauses, especially for familiar events. This shows the use of different syntactic structures can influence conceptual event segmentation even at this most basic level.

Chapter 6 asked whether people remember events described by SVCs as single units. It focused on the taking and putting domain, since Avatime describes many, but not all, putting actions with take-put SVCs which include the preceding taking action in their representation of the putting action. If the taking and putting actions described in an Avatime take-put SVC are remembered as a single event, then seeing one action should make Avatime speakers more likely to falsely recognise the other. To investigate this, SVC and non-SVC taking and putting action pairs were filmed and Avatime-, Dutch- and English-speaking participants were shown one action from each pair before performing a recognition memory task on all actions. The prediction was that Avatime speakers would show greater accuracy for previously unseen non-SVC than SVC actions, while speakers of the other languages would show no difference between SVC and non-SVC actions. However, speakers of all three languages showed greater accuracy for both previously seen and new non-SVC actions. This points towards a commonality in how speakers of all three languages attend to these actions, although the events are only linguistically distinguished in Avatime. There was no evidence for Avatime speakers remembering the SVC action pairs as single events.

## 7.2 Theoretical implications

So, in the end, what can we conclude from these studies? This thesis considered a general question regarding the relations between linguistic and conceptual event discretization and within that the specific question of whether Avatime SVCs refer to single events. Let me first discuss the specific before moving to the more general.

### 7.2.1 SVCs and single events

The results of this thesis suggest that, yes, SVCs do in fact refer to single events. This was most directly shown in the co-speech gesture study of Chapter 4, where SVCs were found to occur with single rather than multiple gestures. This suggests the speakers selected single event representations for the purposes of these utterances, rather than focusing on a more fine-grained representation with distinct subevents. Using an SVC may then be one of the ways in which the speakers communicate this single event framing to their addressees. Notably, this single event framing holds

even though the event information is explicitly distributed throughout the clause with the individual verbs within the SVC describing different aspects or even subevents of the whole.

The other studies tested the reach of this association between SVCs and single events beyond the context of thinking-for-speaking. In Chapter 5, priming with SVCs rather than coordinated clauses was associated with more holistic event discretization during perception. This suggests the use of SVCs may influence the way people perceive events. While the memory study in Chapter 6 was inconclusive, the influence on perception combined with the well-documented influences of segmentation during perception on memory (e.g. Kurby & Zacks 2008; Sargent et al. 2013; Zacks et al. 2007) suggests events described with SVCs are very likely also remembered more holistically. These studies thus suggest the connection between SVCs and single events may come to extend beyond the particular utterance.

Notably, the connection between SVCs and single events was found to hold for the sequential SVCs as well as the nuclear and core types. The nuclear and core SVCs are more intuitively related to single events, since the individual verbs tend to express simultaneously occurring aspects of an event, e.g. posture and activity as in *odi nu* 'he sat listening', rather than temporally distinct subevents. Some researchers (e.g. Bohnemeyer et al. 2007; Pawley 2011) have also suggested that SVCs such as the Avatime sequential type, which allow independent modification by temporal adverbials, would actually refer to multiple events in contrast to other types of SVCs. The Avatime sequential SVCs were examined in two studies in this thesis and in both cases the results suggest a connection with a single event representation. Firstly, while there were only a few sequential SVCs in the data used for the co-speech gesture study in Chapter 4, there were no apparent differences between the SVC subtypes. Secondly, and more convincingly, the study of segmentation during perception in Chapter 5 showed an influence of priming with SVCs versus coordinate clauses, even though the only SVCs used were in fact of the sequential type. This suggests that even these types of SVCs involve a more unified event representation.

Although the present thesis has focused on SVCs in Avatime, the findings are likely to extend to other serializing languages. The association between SVCs and single events has been commonly claimed for serializing languages in general (e.g. Aikhenvald 2006; Bisang 2009; Comrie 1995). Given the similarities between SVCs

in Avatime and those found in other languages, both in terms of their morphosyntactic properties and the functions for which they are employed, I would expect the connection to extend to most SVCs in most serializing languages. Some subtypes of SVCs with a looser connection between their verbs, such as the narrative SVCs of Kalam discussed by Pawley (1987; 2008; 2011), may be more likely to diverge from the association with single events. The extent and nature of the variability or consistency of the relation between SVCs and single events both within and across individual languages would be a promising area for future research as it would help illuminate the motivations and nature of the relationship between these syntactic constructions and conceptual event representations.

### 7.2.2 Relations between conceptual and linguistic event units

Taking a broader view, what do these findings tell us about the way we discretize activity, and the alignments between conceptual and linguistic event units more generally?

Firstly, the study in Chapter 4 shows a strong alignment between linguistic structure and event structure within the conceptual message to be communicated. All of the single clause constructions – SVCs, non-finite subordinate constructions, and simple monoverbal clauses – occurred almost exclusively with single event gestures. In contrast, all multiclausal constructions occurred regularly with multiple event-related gestures and seldom with single totally overlapping gestures. This aligns with previous suggestions that conceptual message formulation aligns with the structural unit of the clause (e.g. Bock 1982; Garrett 1982; Levelt 1989) and that the clausal unit is the one of most relevance for conceptual event structure (e.g. Evans 2010; Jackendoff 1991). In contrast, it appears to go against recent suggestions by Bohnemeyer and colleagues (Bohnemeyer et al. 2007; Bohnemeyer et al. 2011; Bohnemeyer & Van Valin 2009) that there is a sub-clausal unit – the core from Role and Reference Grammar – which aligns more closely with the conceptual event structure.

Outside of thinking-for-speaking, the ways people segmented events were more varied and flexible, but still showed the potential of influence from linguistic structure. In Chapter 5, priming with SVCs versus coordinate clauses influenced how Avatime speakers segmented events during perception. This influence of linguistic structure was greater for more familiar events and likely mediated through event schemata. As people gain understandings of various event types in their

communities and the ways these events are described they form event scripts or schemata for them (e.g. Schank & Abelson 1977). These event schemata would include hierarchical event information and provide a range of possible segmentation patterns. Using SVCs, or other linguistic constructions, may filter or emphasize some of these ways of framing the event. This would then have a follow on effect for the specific model a person generates for the event they are currently experiencing and so influence the way they segment it (Zacks et al. 2007). This influence does not, however, lead to a strong alignment between linguistic and conceptual structures beyond thinking-for-speaking. There was still a great deal of variation in where the participants in the priming study placed their event boundaries. The studies in Chapter 6 also showed no strong tendencies to remember activities described by SVCs as single units. Thus, while linguistic structures may have some influence on the ways events are segmented beyond thinking-for-speaking, they do not appear to strongly constrain these representations.

It should be noted that the participants in the thinking-for-speaking study of Chapter 4 were considerably older than the teenagers who participated in the other studies. One might then wonder whether age is driving the differences found between the studies. For instance, Lucy and Gaskins (2001) noted that influences of language on thought developed with age and could lag behind acquisition of linguistic structures. This is not a very likely explanation for the present results, however. The main issue here is that the younger participants showed greater variability in their segmentation patterns than the older participants. However studies comparing event segmentation patterns in older and younger adults have found that older adults, such as the participants in the co-speech gesture study, typically segment more variably than young adults of around 20 years of age (Zacks, Speer, et al. 2006). In contrast, greater alignment between linguistic constructions and conceptual event units is exactly what one would predict in the thinking-for-speaking context. Thus, it is much more likely that the differences found here reflect real differences in the alignment between linguistic and conceptual event units across different domains of thinking.

The findings of this thesis therefore support the view that conceptualizations start out as rich and varied, and only become constrained and aligned with linguistic structures as part of the process of speech formulation (e.g. Levelt 1989; Slobin 1987; Slobin 1996a). This aligning function of speech is also in line with

previous findings in research on event segmentation (Zacks, Tversky & Iyer 2001), which suggests the alignment between segmentations at different granularities and times increases when people describe the events as they segment them. Thus linguistic encoding appears to provide a scaffold for event segmentation which fixes the particular event structure chosen during the description.

This linguistic scaffold does not, however, completely constrain the conceptual event structure. For instance, in the co-speech gesture study of Chapter 4, several instances were noted where single clauses occurred with multiple event gestures, and vice-versa where multiclause constructions occurred with single overlapping gestures. This result agrees with work showing that multiple representations can be active with one appearing in gesture and another in speech (e.g. Goldin-Meadow 2003; Pine, Lufkin & Messer 2004). Both a multiple and a single event representation could then still be active at the point where the articulation of the spoken and gestural components diverge. So while conceptual and linguistic event units are generally very closely aligned at this stage of thinking-for-speaking, they are not completely isomorphic and remain separate systems which may deviate from each other.

The studies described in this thesis suggest a deep and multifaceted connection between conceptual and linguistic event segmentation. In the act of describing an event, the conceptual representation becomes fixed and there is a close alignment between the chosen conceptual and linguistic structures. More generally, a lifetime of communicating with others about events, and the routine experience of hearing and producing event descriptions, informs the event schemata people construct and influences the selection of these schemata in particular contexts. These studies thus largely confirm the suggestion made by Pawley while working on SVCs in Kalam.

> "What is reported as an event or happening is the outcome of a complex interaction between many variables: between certain physical phenomena and the speaker's sensory impressions of these phenomena; habits and expectations, and limitations of attention span, biasing the speaker's interpretation of these impressions; limitations of long-term memory influencing his recall; and, among other things, structural patterns provided by the language which shape the form of his report and possibly his initial perception and memory of what happened." (Pawley 1987: 330)

To sum up, this thesis has investigated the relations between Avatime SVCs and conceptual event units during thinking-for-speaking, perception, and memory. The studies of co-speech gestures and segmentation during perception provided converging experimental evidence supporting the oft-claimed connection between SVCs and single events. Furthermore, the studies in this thesis have shown connections between linguistic and conceptual event units, with a close alignment between conceptual units and clauses during thinking-for-speaking, as well as evidence of influences of linguistic structure on conceptual event segmentation more generally.

## References


Aboh, Enoch O. 2009. Clause Structure and Verb Series. *Linguistic Inquiry* 40(1). 1–33.

Aboh, Enoch O. & James Essegbey. 2010. Introduction. In Enoch O. Aboh & James Essegbey (eds.), *Topics in Kwa Syntax*, xi–xv. Dordrecht: Springer.

Aikhenvald, Alexandra Y. 2006. Serial verb constructions in typological perspective. In Alexandra Y. Aikhenvald & R. M. W. Dixon (eds.), *Serial verb constructions: A cross-linguistic typology*, 1–68. Oxford: Oxford University Press.

Alferink, Inge & Marianne Gullberg. 2014. French-Dutch bilinguals do not maintain obligatory semantic distinctions: Evidence from placement verbs. *Bilingualism: Language and Cognition* 17(1). 22–37. doi:10.1017/S136672891300028X.

Ameka, Felix K. 1995. The linguistic construction of space in Ewe. *Cognitive Linguistics* 6(2/3). 139–182.

Ameka, Felix K. 2003. Multiverb constructions in a West African areal typological perspective. *Proceedings of TROSS 3*. http://edvarda.hf.ntnu.no/ling/tross/TROSS03-toc.html.

Ameka, Felix K. 2006a. Ewe serial verb constructions in their grammatical context. In Alexandra Y. Aikhenvald & R. M. Dixon (eds.), *Serial verb constructions: A cross-linguistic typology*, 124–143. Oxford: Oxford University Press.

Ameka, Felix K. 2006b. Grammars in contact in the Volta Basin (West Africa): On contact induced grammatical change in Likpe. In Alexandra Y. Aikhenvald & R. M. W. Dixon (eds.), *Grammars in contact: A crosslinguistic typology*, 114–142. Oxford: Oxford University Press.

Ameka, Felix K. 2008. Aspect and modality in Ewe: A survey. In Felix K. Ameka & Mary Esther Kropp Dakubu (eds.), *Aspect and modality in Kwa languages*, 135–194. Amsterdam: Benjamins.

Ameka, Felix K. 2009. Likpe. In Gerrit J. Dimmendaal (ed.), *Coding participant marking: Construction types in twelve African languages*, 239–280. Amsterdam: Benjamins.

Ameka, Felix K. 2010. Information packaging constructions in Kwa: Microvariation and typology. In Enoch O. Aboh & James Essegbey (eds.), *Topics in Kwa syntax (Studies in natural language and linguistic theory 78)*, 141–176. Dordrecht: Springer.

Avrahami, Judith & Yaakov Kareev. 1994. The emergence of events. *Cognition* 53. 239–261.

Bach, Emmon. 1986. The algebra of events. *Linguistics and Philosophy* 9(1). 5–16.

Baird, Jodie A. & Dare Baldwin. 2001. Making sense of human behavior: Action parsing and intentional inference. In B. F. Malle, L. J. Moses & Dare Baldwin (eds.), *Intentions and intentionality: Foundations of social cognition*, 193–206. Cambridge, MA: MIT Press.

Baker, Brett & Mark Harvey. 2010. Complex predicate formation. In Mengistu Amberber, Brett Baker & Mark Harvey (eds.), *Complex predicates: Cross-linguistic perspectives on event structure*, 13–47. Cambridge: Cambridge University Press.

Baker, Mark C. 1989. Object sharing and projection in serial verb constructions. *Linguistic Inquiry* 20(4). 513–553. doi:http://www.jstor.org/stable/4178644.

Bamgbose, Ayo. 1974. On serial verbs and verbal status. *Journal of West African Languages* 9(1). 17–48.

Barr, Dale J., Roger Levy, Christoph Scheepers & Harry J. Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68. 255–278. doi:10.1016/j.jml.2012.11.001.

Barsalou, Lawrence W. 1992. Frames, concepts, and conceptual fields. In Adrienne Lehrer & Eva Feder Kittay (eds.), *Frames, fields, and contrasts: New essays in semantic and lexical organization*, 21–74. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Bates, Douglas, Reinhold Kliegl, Shravan Vasishth & Harald Baayen. 2015. Parsimonious mixed models. *arXiv*. http://arxiv.org/abs/1506.04967v1.

Bates, Douglas, Martin Maechler, Ben Bolker & Steven Walker. 2015. *lme4: Linear mixed-effects models using Eigen and S4*. http://CRAN.R-project.org/package = lme4.

Beavers, John, Beth Levin & Shiao Wei Tham. 2010. The typology of motion expressions revisited. *Journal of Linguistics* 46. 331–377.

Bickel, Balthasar. 2010. Capturing particulars and universals in clause linkage: a multivariate analysis. In Isabelle Bril (ed.), *Clause linking and clause hierarchy: syntax and pragmatics*. Amsterdam/Philadelphia: John Benjamins.

Bisang, Walter. 2009. Serial verb constructions. *Language and Linguistics Compass* 3(3). 792–814. doi:10.1111/j.1749-818X.2009.00128.x.

Blench, Roger. 2009. Do the Ghana-Togo mountain languages constitute a genetic group? *Journal of West African Languages* 36(1&2). 19–36.

Bloom, Paul & Frank C. Keil. 2001. Thinking through language. *Mind and Language* 16(4). 351–367.

Bobuafor, Mercy. 2013. A grammar of Tafi. Leiden: Leiden University PhD dissertation.

Bock, Kathryn. 1982. Towards a cognitive psychology of syntax: Information processing contributions to sentence formulation. *Psychological Review* 89. 1–47.

Bock, Kathryn. 1986. Syntactic persistence in language production. *Cognitive Psychology* 18. 355–387.

Bock, Kathryn & Zenzi M. Griffin. 2000. The Persistence of Structural Priming: Transient Activation or Implicit Learning? *Journal of Experimental Psychology: General* 129(2). 177–192.

Bohnemeyer, Jürgen, Sonja Eisenbeiss & Bhuvana Narasimhan. 2006. Ways to go: Methodological considerations in Whorfian studies on motion events. *Essex Research Reports in Linguistics* 50. 1–20.

Bohnemeyer, Jürgen, Nicholas J. Enfield, James Essegbey, Iraide Ibarretxe-Antuñano, Sotaro Kita, Friederike Lüpke & Felix K. Ameka. 2007. Principles of event segmentation in language: The case of motion events. *Language* 83(3). 495–532. doi:http://www.jstor.org/stable/40070900.

Bohnemeyer, Jürgen, Nick Enfield, James Essegbey & Sotaro Kita. 2011. The Macro-Event property: The segmentation of causal chains. In Jürgen Bohnemeyer & Eric Pederson (eds.), *Event representation in language and cognition*, 43–67. (Language Culture and Cognition 11). Cambridge: Cambridge University Press.

Bohnemeyer, Jürgen & Eric Pederson (eds.). 2011. *Event representation in language and cognition*. (Language Culture and Cognition 11). Cambridge: Cambridge University Press.

Bohnemeyer, Jürgen & Robert Van Valin. 2009. The macro-event property and the LSC. University of California, Berkeley. http://www.acsu.buffalo.edu/~jb77/macro=core_JB&Van_handout.pdf.

Boroditsky, Lera. 2011. How languages construct time. In Stanislas Dehaene & Elizabeth Brannon (eds.), *Space, time and number in the brain: Searching for the foundations of mathematical thought*, 333–341. Elsevier.

Boroditsky, Lera, Lauren A. Schmidt & Webb Phillips. 2003. Sex, syntax, and semantics. In Dedre Gentner & Susan Goldin-Meadow (eds.), *Language in mind:*

*advances in the study of language and thought*, 61–79. Cambridge, MA: MIT Press.

Boster, James S. & Jeffrey C. Johnson. 1989. Form or function: A comparison of expert and novice judgements of similarity among fish. *American Anthropologist* 91(4). 866–889.

Bowden, John. 2008. Verb serialisation in Taba. In Gunter Senft (ed.), *Serial verb constructions in Austronesian and Papuan languages*, 75–98. Canberra: Pacific Linguistics.

Bowerman, Melissa & Penelope Brown (eds.). 2008. *Crosslinguistic perspectives on argument structure: Implications for learnability*. New York: Lawrence Erlbaum Associates.

Bowerman, Melissa, Marianne Gullberg, Asifa Majid & Bhuvana Narasimhan. 2004. Put project: the cross-linguistic encoding of placement events. In Asifa Majid (ed.), *Field manual Volume 9*, 10–24. Nijmegen: Max Planck Institute for Psycholinguistics. http://fieldmanuals.mpi.nl/volumes/2004/put-project/.

Bril, Isabelle. 2004. Complex nuclei in Oceanic languages: Contribution to an areal typology. In Isabelle Bril & Françoise Ozanne-Rivierre (eds.), *Complex Predicates in Oceanic Language: Studies in the Dynamics of Binding and Boundness*, 1–48. Berlin: Mouton de Gruyter.

Brown, Dunstan, Marina Chumakina & Greville G. Corbett (eds.). 2012. *Canonical morphology and syntax*. Oxford: Oxford University Press.

Bruce, Les. 1988. Serialisation: from syntax to lexicon. *Studies in Language* 12. 19–49. doi:10.1075/sl.12.1.03bru.

Brydon, Lynne. 1981. Rice, yams and chiefs in Avatime: speculations on the development of a social order. *Africa* 51(2). 659–677.

Brydon, Lynne. 1996. Women chiefs and power in the Volta region of Ghana. *Journal of Legal Pluralism and Unoffical Law* 37–38. 227–247.

Brydon, Lynne. 2008. Constructing Avatime: questions of history and identity in a West African polity, c. 1690s to the twentieth century. *Journal of African History* 49. 23–42.

Butterworth, Brian & Geoffrey Beattie. 1978. Gesture and silence as indicators of planning in speech. In R. N. Campbell & P. T. Smith (eds.), *Recent advances in the psychology of language 4: Formal and experimental approaches*, 347–360. London: Plenum.

Casasanto, Daniel. 2013. Gesture and language processing. *Encyclopedia of the mind*, 372–374. Oaks, CA: Sage Publications.

Casasanto, Daniel & Kyle Jasmin. 2012. The Hands of Time: Temporal gestures in English speakers. *Cognitive Linguistics* 23(4). 643–674.

Casati, Roberto & Achille C. Varzi (eds.). 1996. *Events*. Aldershot, England: Dartmouth.

Casati, Roberto & Achille C. Varzi. 2008. Event Concepts. In Thomas F. Shipley & Jeffrey M. Zacks (eds.), *Understanding events: from perception to action*, 31–53. Oxford: Oxford University Press.

Chang, Franklin, Gary S. Dell & Kathryn Bock. 2006. Becoming syntactic. *Psychological Review* 113(2). 234–272.

Choi, Soonja. 2006. Influence of language-specific input on spatial cognition: categories of containment. *First Language* 26(2). 207–232.

Christaller, Johann Gottlieb. 1875. *A grammar of the Asante and Fante language called Tshi (Chwee, Twi)*. Basel: Basel German Evangelical Mission. Reprinted in 1964 by Gregg Press Inc., Ridgewood, N.J.

Cleland, Carol. 1991. On the individuation of events. *Synthese* 86. 229–254.

Cohen, Claudia E. & Ebbe B. Ebbesen. 1979. Observational goals and schema activation: A theoretical framework for behavior perception. *Journal of Experimental Social Psychology* 15. 305–329.

Collins, Chris. 1997. Argument sharing in serial verb constructions. *Linguistic Inquiry* 28(3). 461–497.

Comrie, Bernard. 1995. Serial verbs in Haruai (Papua New Guinea) and their theoretical implications. In J. Bouscaren, J. Franckel & S. Robert (eds.), *Langues et langage: Problèmes et raisonnement en linguistique, mélanges offerts à Antoine Culioli*, 25–37. Paris: University Presses of France.

Croft, William. 1990. Possible verbs and the structure of events. In S. L. Tsohatzidis (ed.), *Meanings and prototypes: Studies in linguistic categorization*, 48–73. London: Routledge.

Croft, William, Jóhanna Barðdal, Willem B. Hollmann, Violeta Sotirova & Chiaki Taoka. 2010. Revising Talmy's typological classification of complex event constructions. In Hans C. Boas (ed.), *Contrastive studies in construction grammar*, 201–236. Amsterdam/Philadelphia: John Benjamins.

Crowley, Terry T. 2002. *Serial verbs in Oceanic: a descriptive typology*. Oxford: Oxford University Press.

Dahl, Östen. 2015. How WEIRD are WALS languages? MPI-EVA, Leipzig.

Davidson, Donald. 1969. The individuation of events. In Roberto Casati & Achille C. Varzi (eds.), *Events*, 265–284. Aldershot, England: Dartmouth. (Reprinted from Essays in honor of Carl G. Hempel, pp. 216-234, by N. Rescher, (Ed.), 1969, Dordrecht: D. Reidel).

Déchaine, Rose-Marie. 1993. Serial verb constructions. In Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld & Theo Vennemann (eds.), *Syntax: An international handbook of contempory research*, vol. 1, 799–825. Berlin, New York: Walter de Gruyter.

Defina, Rebecca. in press. Tense, aspect and mood in Avatime. *Afrika und Uberzee*.

Defina, Rebecca. 2009. Aspect and modality in Avatime. Leiden: Leiden University MPhil Dissertation.

Defina, Rebecca. 2016. Serial verb constructions and their subtypes in Avatime. *Studies in Language*.

Diller, A. V. N. 2006. Thai serial verbs: Cohesion and culture. In Alexandra Y. Aikhenvald & R. M. W. Dixon (eds.), *Serial verb constructions: A cross-linguistic typology*, 160–177. Oxford: Oxford University Press.

Dimmendaal, Gerrit J. 2001. Areal diffusion versus genetic inheritance: an African perspective. In Alexandra Y. Aikhenvald & R. M. W. Dixon (eds.), *Areal diffusion and genetic inheritance*, 358–392. Oxford: Oxford University Press.

Dingemanse, Mark. 2011. The meaning and use of ideophones in Siwu. Nijmegen: Radboud University Nijmegen Ph.D. Thesis.

Dixon, R. M. W. 2006. Serial verb constructions: conspectus and coda. *Serial Verb Constructions: A Crosslinguistic Typology*, 338–350. New York: Oxford University Press.

Dolscheid, Sarah, Shakila Shayan, Asifa Majid & Daniel Casasanto. 2013. The thickness of musical pitch: Psychophysical evidence for linguistic relativity. *Psychological Science* 24. 613–621. doi:10.1177/0956797612457374.

Dorvlo, Kofi. 2008. A grammar of Logba. Leiden: University of Leiden Ph.D. Thesis.

Dougherty, Janet W. & Charles M. Keller. 1982. Taskonomy: a practical approach to knowledge structures. *American Ethnologist* 9(4). 763–774.

DuBrow, Sarah & Lila Davachi. 2013. The influence of context boundaries on memory for the sequential order of events. *Journal of Experimental Psychology: General* 142(4). 1277–1286.

Durie, Mark. 1997. Grammatical structures in verb serialization. In Alex Alsina, Joan Bresnan & Peter Sells (eds.), *Complex predicates*, 289–354. Stanford, California: CSLI Publictaions.

Enfield, Nicholas J. 2002. Cultural logic and syntactic productivity: Associated posture constructions in Lao. In Nicholas J. Enfield (ed.), *Ethnosyntax: Explorations in culture and grammar*, 231–258. Oxford: Oxford University Press.

Enfield, Nicholas J. 2007. *A grammar of Lao*. Berlin: Mouton de Gruyter.

Enfield, Nicholas J. 2009. *The anatomy of meaning: Speech, gesture, and composite utterances*. Cambridge: Cambridge University Press.

Esteve-Gibert, Núria & Pilar Prieto. 2013. Prosodic structure shapes the temporal realization of intonation and manual gesture movements. *Journal of Speech, Language, and Hearing Research* 56. 850–864. doi:10.1044/1092-4388(2012/12-0049).

Evans, Nicholas. 2010. Complex events, propositional overlay, and the special status of reciprocal clauses. In Sally Rice & John Newman (eds.), *Empirical and experimental methods in cognitive/functional research*. Stanford: CSLI Publictaions.

Fausey, Caitlin M. & Lera Boroditsky. 2011. Who dunnit? Cross-linguistic differences in eye-witness memory. *Psychonomic Bulletin & Review* 18(1). 150–157. doi:10.3758/s13423-010-0021-5.

Fausey, Caitlin M., Bria L. Long, Aya Inamori & Lera Boroditsky. 2010. Constructing agency: the role of language. *Frontiers in Psychology* 1(162).

Filipović, Luna. 2011. Speaking and remembering in one or two languages: bilingual vs. monolingual lexicalisation and memeory for motion events. *International Journal of Bilingualism* 15(4). 466–485.

Fillmore, Charles. 1977. Scenes-and-frames semantics. In Antonio Zambolli (ed.), *Linguistic structure processing*, 55–82. Amsterdam: North-Holland Publishing Company.

Finkbeiner, Matthew, Janet Nicol, Delia Greth & Kumiko Nakamura. 2002. The role of language in memory for actions. *Journal of Psycholinguistic Research* 31(5). 447–457.

Foley, William A. 1997. Polysynthesis and complex verbs formation: The case of applicatives in Yimas. In Alex Alsina, Joan Bresnan & Peter Sells (eds.), *Complex predicates*, 355–395. Stanford, California: Center for the Study of Language and Information, Stanford University.

Foley, William A. 2010a. Events and serial verb constructions. In Brett Baker & Mark Harvey (eds.), *Complex predicates: Cross-linguistic perspectives on event structure*, 79–109. Cambridge: Cambridge University Press.

Foley, William A. 2010b. Clause linkage and nexus in Papuan languages. In Isabelle Bril (ed.), *Clause linking and clause hierarchy: syntax and pragmatics*, 27–50. Amsterdam/Philadelphia: John Benjamins.

Foley, William A. & Mike Olson. 1985. Clausehood and verb serialization. In Johanna Nichols & Anthony C. Woodbury (eds.), *Grammar inside and outside the clause*, 17–60. Cambridge: Cambridge University Press.

Foley, William A. & Robert Van Valin. 1984. *Functional syntax and universal grammar*. Cambridge: Cambridge University Press.

Folli, Raffaela & Heidi Harley. 2006. What language says about the psychology of events. *Trends in Cognitive Sciences* 10(3). 91–92.

Ford, Kevin C. 1971a. Aspects of Avatime syntax. Legon: University of Ghana PhD.

Ford, Kevin C. 1971b. Noun classes, concord and markedness in Avatime. *Annales de L'Université d'Abidjan* Ser. H. 141–149.

Frank, Michael C., Daniel L. Everett, Evelina Fedorenko & Edward Gibson. 2008. Number as a cognitive technology: Evidence from Pirahã language and cognition. *Cognition* 108(3). 819–824. doi:10.1016/j.cognition.2008.04.007.

Funke, Emil. 1909. Versuch einer Grammatik der Avatimesprache. *Mitteilungen des Seminars für orientalische Sprachen* 12(3). 287–336.

Garber, Philip & Susan Goldin-Meadow. 2002. Gesture offers insight into problem-solving in adults and children. *Cognitive Science* 26. 817–831. doi:10.1207/s15516709cog2606_5.

Garrett, Merrill F. 1982. Production of speech: Observations from normal and pathological language use. In Andrew W. Ellis (ed.), *Normality and pathology in cognitive functions*, 19–76. London: Academic Press.

Gennari, Silvia P., Steven A. Sloman, Barbara C. Malt & W. Tecumseh Fitch. 2002. Motion events in language and cognition. *Cognition* 83. 49–79.

Gentner, Dedre. 1982. Why nouns are learned before verbs: linguistic relativity versus natural partitioning. In S. A. Kuczaj (ed.), *Language Development*, vol. 2: Language, Thought, and Culture, 301–344. Hillsdale, New Jersey: Lawrence Erlbaum.

Gentner, Dedre. 2006. Why verbs are hard to learn. In K. Hirsh-Pasek & R. Golinkoff (eds.), *Action meets word: How children learn verbs*, 544–564. Oxford: Oxford University Press.

Gentner, Dedre & Lera Boroditsky. 2001. Individuation, relativity and early word learning. In Melissa Bowerman & Stephen C. Levinson (eds.), *Language acquisition and conceptial development*, 215–256. Cambridge: Cambridge University Press.

Gerrig, Richard J. & Mahzarin R. Banaji. 1994. Language and thought. *Thinking and problem solving: Handbook of perception and cognition*, 233–261. San Diego, CA: Academic Press.

Gilbert, Aubrey L., Terry Regier & Paul Kay. 2006. Whorf hypothesis is supported in the right visual field but not the left. *Proceedings of the National Academy of Sciences of the USA* 103(2). 489–494.

Givón, Talmy. 1991. Serial verbs and event cognition in Kalam: an empirical study of cultural relativity. In Clare Lefebvre (ed.), *Serial verbs: grammatical, comparative and universal grammar*, 137–184. Amsterdam: John Benjamins.

Goldin-Meadow, Susan. 2003. *Hearing gesture: How our hands help us think*. Cambridge, MA: Harvard University Press.

Goldin-Meadow, Susan. 2006. Talking and thinking with our hands. *Current Directions in Psychological Science* 15. 34–39.

Goldin-Meadow, Susan & Sian L. Beilock. 2010. Action's influence on thought: The case of gesture. *Perspectives on Psychological Science* 5. 664–674.

Goldman, Alvin I. 1971. The individuation of action. *Journal of Philosophy* 68. 761–774.

Goldman, Alvin I. 2007. A Program for "Naturalizing" Metaphysics, with Application to the Ontology of Events. *The Monist* 90(3). 457–479.

Goldman-Eisler, Frieda. 1968. *Psycholinguistics: experiments in spontaneous speech*. New York: Academic Press.

Gullberg, Marianne. 2011. Language-specific encoding of placement events in gestures. In Jürgen Bohnemeyer & Eric Pederson (eds.), *Event representation in language and cognition*, 166–188. Cambridge: Cambridge University Press.

Haberlandt, Karl & Arthur C. Graesser. 1989. Processing of new arguments at clause boundaries. *Memory & Cognition* 17(2). 186–193.

Hale, Ken & Jay Keyser. 1993. On argument structure and the lexical expression of syntactic relations. In Ken Hale & Jay Keyser (eds.), *The view from Building 20*, 53–109. Cambridge: Mass.: The MIT Press.

Hard, Bridgette M., Gabriel Recchia & Barbara Tversky. 2011. The shape of action. *Journal of Experimental Psychology: General* 140(4). 586–604. doi:10.1037/a0024310.

Hard, Bridgette M., Barbara Tversky & David S. Lang. 2006. Making sense of abstract events: Building event schemas. *Memory & Cognition* 34(6). 1221–1235.

Haspelmath, Martin. 2016. The serial verb construction: Comparative concept and cross-linguistic generalizations. *Language and Linguistics* 17(3). 291–319.

Heider, Fritz. 1958. *The psychology of interpersonal relations*. New York: Wiley.

Heine, Bernd. 1968. *Die Verbreitung und Gliederung der Togorestsprachen*. Berlin: Reimer.

Heine, Bernd. 2008. From “Togorest” to GTM: Some reflections. Ho, Ghana.

Hellan, Lars, Dorothee Beermann & Eli Sætherø Andenes. 2003. Towards a typology of serial verb constructions in Akan. In Mary Esther Kropp Dakubu & Emmanuel Kweku Osam (eds.), *Studies in the languages of the Volta basin*, vol. 1, 87–110. Legon, Ghana: Linguistics Department, University of Ghana.

Henrich, Joseph, Steven J. Heine & Ara Norenzayan. 2010. The weirdest people in the world. *Behavioral and Brain Sciences* 33(2–3). 61–83. doi:http://dx.doi.org/10.1017/S0140525X0999152X.

Himmelmann, Nikolaus P. 2013. Some observations on clausehood and event segmentation in Austronesian and Papuan Languages. Paper presented at the Sixth Austronesian and Papuan Languages and Linguistics conference, School of African and Asian Studies, London.

Hübler, Axel. 2007. *The nonverbal shift in Early Modern English conversation*. Amsterdam: John Benjamins.

Huff, Markus & Stephan Schwan. 2008. Verbalising events:overshadowing or facilitation? *Memory & Cognition* 36(2). 392–402. doi:10.3758/MC.36.2.392.

IBM Corp. 2011. *IBM SPSS Statistics for Windows*. Armonk, NY: IBM Corp.

Jackendoff, Ray. 1990. *Semantic structures*. Cambridge: Mass.: MIT Press.

Jackendoff, Ray. 1991. Parts and boundaries. In Beth Levin & Steven Pinker (eds.), *Lexical and conceptual semantics*, 9–46. Amsterdam: Elsevier Science Publishers.

Jarkey, Nerida. 1991. Serial verb constructions in White Hmong: a functional approach. Sydney: University of Sydney PhD Dissertation.

Just, Marcel Adam, Patricia A. Carpenter & Jacqueline D. Woolley. 1982. Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General* 111(2). 228–238.

Kendon, Adam. 1972. Some relationships between body motion and speech. In A. Seigman & B. Pope (eds.), *Studies in dyadic communication*, 177–216. Elmsford, New York: Pergamon Press.

Kendon, Adam. 1980. Gesticulation and speech: two aspects of the process of utterance. In Mary R. Key (ed.), *The relationship of verbal and nonverbal communication*, 207–228. The Hague: Mouton.

Kendon, Adam. 1986. Some reasons for studying gesture. *Semiotica* 62(1/2). 3–28. doi:10.1515/semi.1986.62.1-2.3.

Kendon, Adam. 2004. *Gesture: visible action as utterance*. Cambridge: Cambridge University Press.

Kenny, Anthony. 1963. *Actions, emotion, and will*. Humanities Press.

Kersten, Alan W., Christian A. Meissner, Julia Lechuga, Bennett L. Schwartz, Justin S. Albrechtsen & Adam Iglesias. 2010. English Speakers Attend More Strongly Than Spanish Speakers to Manner of Motion When Classifying Novel Objects and Events. *Journal of Experimental Psychology: General* 139(4). 638–653.

Kießling, Roland. 2011. *Verbal serialisation in Isu (West-Ring) - a Grassfields language of Cameroon*. Cologne, Germany: Rüdiger Köppe Verlag.

Kim, Jaegwon. 1976. Events as property exemplifications. In Roberto Casati & Achille C. Varzi (eds.), *Events*, 117–136. Aldershot, England: Dartmouth. (Reprinted from Action theory, pp. 159-177, by Myles Brand and Douglas Walton, (Eds.), 1976, Dordrecht: D. Reidel).

Kita, Sotaro. 1990. The temporal relationship between gesture and speech: A study of Japanese-English bilinguals. Chicago: University of Chicago Master's thesis.

Kita, Sotaro & Asli Özyürek. 2003. What does cross-linguistic variation in semantic coordination of speech and gesture reveal?: Evidence for and interface representation of spatial thinking and speaking. *Journal of Memory and Language* 48. 16–32. doi:10.1016/S0749-596X(02)00505-3.

Kita, Sotaro, Asli Özyürek, Shanley Allen, Amanda Brown, Reyhan Furman & Tomoko Ishizuka. 2007. Relations between syntactic encoding and co-speech gestures: Implications for a model of speech and gesture production. *Language and Cognitive Processes* 22(8). 1212–1236. doi:10.1080/01690960701461426.

Komsta, Lukasz & Frederick Novomestky. 2012. *Moments package*.

Krauss, Robert M., Yihsiu Chen & Purnima Chawla. 1996. Nonverbal behavior and nonverbal communication: What do conversational hand gestures tell us? In M. Zanna (ed.), *Advances in experimental social psychology*, 389–450. San Diego, CA: Academic Press.

Kropp Dakubu, Mary Esther. in press. Sub-classifying the languages of the lower Volta valley: towards redefining Kwa. *STUF - Language Typology and Universals*.

Kropp Dakubu, Mary Esther & Kevin C. Ford. 1988. The central-Togo languages. In Mary Esther Kropp Dakubu (ed.), *The languages of Ghana*, 119–154. London: Kegan Paul International.

Kropp, Mary Esther. 1967. *Lefana, Akpafu and Avatime with English glosses*. (Comparative African Wordlists 3). Legon, Ghana: Institute of African Studies.

Kurby, Christopher A. & Jeffrey M. Zacks. 2008. Segmentation in the perception and memory of events. *Trends in Cognitive Sciences* 12(2). 72–79.

Ladd, D. Robert. 1996. *Intonational phonology*. Cambridge: Cambridge University Press.

Landis, J. Richard & Gary G. Koch. 1977. The measurement of observer agreement for categorical data. *Biometrics* 33. 159–174.

Lassiter, G. Daniel & R. David Slaw. 1991. The unitization and memory of events. *Journal of Experimental Psychology: General* 120(1). 80–82. doi:10.1037/0096-3445.120.1.80.

Le Guen, Olivier, Jana Samland, Thomas Friedrich, Daniel Hanus & Penelope Brown. 2015. Making sense of (exceptional) causal relations: A cross-cultural and cross-linguistic study. *Frontiers in Psychology* 6(1645). doi:10.3389/fpsyg.2015.01645.

Levelt, Willem J.M. 1989. *Speaking*. Cambridge, MA: MIT Press.

Levinson, Stephen C. 2012. Foreword. In J.B. Carroll, Stephen C. Levinson & Penny Lee (eds.), *Language, thought, and reality: selected writings of Benjamin Lee Whorf*, vii–xxiii. 2nd ed. Cambridge, MA: MIT Press.

Lewis, M. Paul, Gary F. Simons & Charles D. Fennig (eds.). 2014. *Ethnologue: Languages of the world, Seventeenth edition*. Dallas, Texas: SIL International. Online version: http://www.ethnologue.com.

Lewis, Marshall. 1993. Verb serialization in Gen and the structure of events. Indiana: Indiana University PhD.

Loehr, Daniel P. 2012. Temporal, structural, and pragmatic synchrony between intonation and gesture. *Laboratory Phonology* 3(1). 71–89. doi:10.1515/lp-2012-0006.

Loftus, Elizabeth F. & John C. Palmer. 1974. Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior* 13. 585–589. doi:10.1016/S0022-5371(74)80011-3.

Lord, Carol. 1993. *Historical change in serial verb constructions*. (Typological Studies in Language 26). Amsterdam: Benjamins.

Loucks, Jeff & Eric Pederson. 2011. Linguistic and non-linguistic categorization of complex motion events. In Jürgen Bohnemeyer & Eric Pederson (eds.), *Event representation in language and cognition*, 108–133. (Language Culture and Cognition 11). Cambridge: Cambridge University Press.

Lucy, John & Suzanne Gaskins. 2001. Grammatical categories and the development of classification preferences: A comparative approach. In Stephen C. Levinson & Melissa Bowerman (eds.), *Language acquisition and conceptual development*, 257–283. Cambridge: Cambridge University Press.

Maddieson, Ian. 1998. Collapsing vowel-harmony and doubly-articulated fricatives: two myths about the Avatime phonological system. In Ian Maddieson & T. J. Hinnebusch (eds.), *Language history and linguistic description in Africa*, 155–166. (Trends in African Linguistics 2). Trenton, NJ/Asmara: Africa World Press.

Maini, Raman & Himanshu Aggarwal. 2009. Study and comparison of various image edge detection techniques. *International Journal of Image Processing* 3(1). 1–11.

Majid, Asifa, James S. Boster & Melissa Bowerman. 2008. The cross-linguistic categorization of everyday events: A study of cutting and breaking. *Cognition* 109. 235–250.

Majid, Asifa, Melissa Bowerman, Sotaro Kita, Daniel B.M. Haun & Stephen C. Levinson. 2004. Can language restructure cognition? The case for space. *Trends in Cognitive Sciences* 8(3). 108–114. doi:10.1016/j.tics.2004.01.003.

Malaia, Evie. 2014. It still isn't over: Event boundaries in language and perception. *Language and Linguistics Compass* 8(3). 89–98. doi:10.1111/lnc3.12071.

Malt, Barbara C. 2012. Mechanisms of relativity: the "language directs attention" hypothesis. Paper presented at the Relations in Relativity, MPI for Psycholinguistics, Nijmegen.

Malt, Barbara C. & Asifa Majid. 2013. How thought is mapped into words. *WIREs Cognitive Science* 4. 583–597. doi:10.1002/wcs.1251.

Malt, Barbara C., Steven A. Sloman, Silvia P. Gennari, Meiyi Shi & Yuan Wang. 1999. Knowing versus naming: similarity and the linguistic categorization of artifacts. *Journal of Memory and Language* 40. 230–262.

Matlock, Teenie, David Sparks, Justin L. Matthews, Jeremy Hunter & Stephanie Huette. 2012. Smashing new results on aspectual framing: How people talk about car accidents. *Studies in Language* 36(3). 699–720. doi:10.1075/sl.36.3.09mat.

Matsumoto, Yo. 2003. Typologies of lexicalization patterns and event integration: clarifications and reformulations. In Masaru Kajita & Shuji Chiba (eds.), *Empirical and theoretical investigations into language: A Festschrift for Masaru Kajita*, 403–418. Tokyo: Kaitakusha.

Mayberry, Rachel & Joselynne Jaques. 2000. Gesture production during stuttered speech: insights into the nature of speech-gesture integration. In David McNeill (ed.), *Language and Gesture*, 199– 215. Cambridge: Cambridge University Press.

McCormack, P. D. & Nancy M. Wright. 1964. The positive skew observed in reaction time distributions. *Canadian Journal of Psychology/Revue canadienne de psychologie* 18(1). 43–51.

McDonough, Laraine, Soonja Choi & Jean M. Mandler. 2003. Understanding spatial relations: flexible infants, lexical adults. *Cognitive Psychology* 46. 229–259.

McNeill, David. 1985. So you think gestures are nonverbal? *Psychological Review* 92(3). 350–371. doi:10.1037/0033-295X.92.3.350.

McNeill, David. 1992. *Hand and mind: What gestures reveal about thought.* Chicago: University of Chicago.

McNeill, David & Susan Duncan. 2000. Growth points in thinking-for-speaking. In David McNeill (ed.), *Language and gesture*, 141–161. Cambridge: Cambridge University Press.

Meyer, Meredith, Dare Baldwin & Kara Sage. 2011. Assessing Young Children's Hierarchical Action Segmentation. In L. Carlson, C. Hölscher & T. Shipley (eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society*, 3156–3161. Austin, TX: Cognitive Science Society.

Michotte, Albert, Georges Thinès & Geneviève Crabbé. 1991. Amodal completion of perceptual structures. In Georges Thinès, Alan Costall & George Butterworth (eds.), *Michotte's experimental phenomenology of perception*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Mooney, Christopher Z. & Robert D. Duval. 1993. *Bootstrapping: a nonparametric approach to statistical inference*. Newbury Park, CA: Sage Publications.

Morrel-Samuels, Palmer & Robert M. Krauss. 1992. Word familiarity predicts temporal asynchrony of hand gestures and speech. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 18. 615–622. doi:10.1037/0278-7393.18.3.615.

Morris, Micael W. & Gregory L. Murphy. 1990. Converging operations on a basic level in event taxonomies. *Memory & Cognition* 18(4). 407–418.

Mourelatos, Alexander P. D. 1978. Events, processes, and states. *Linguistics and Philosophy* 2. 415–434.

Müller, F. Max. 1909. *The science of thought*. London: Longmans, Green & Co.

Narasimhan, Bhuvana, Anetta Kopecka, Melissa F. Bowerman, Marianne Gullberg & Asifa Majid. 2012. Putting and taking events: A crosslinguistic perspective. *Events of putting and taking: a crosslinguistic perspective*, 1–18. Amsterdam/Philadelphia: John Benjamins.

Newton, Ashley M. & Jill G. de Villiers. 2007. Thinking while talking: Adults fail nonverbal false-belief reasoning. *Psychological Science* 18(7). 574–579. doi:10.1111/j.1467-9280.2007.01942.x.

Newtson, Darren. 1973. Attribution and the unit of perception of ongoing behaviour. *Journal of Personality and Social Psychology* 28(1). 28–38.

Newtson, Darren & Gretchen Engquist. 1976. The Perceptual Organization of Ongoing Behavior. *Journal of Experimental Social Psychology* 12. 436–450. doi:10.1016/0022-1031(76)90076-7.

Nobe, Schuichi. 2000. Where do most spontaneous representational gestures actually occur with respect to speech? In David McNeill (ed.), *Language and gesture*, 186–198. Cambridge: Cambridge University Press.

Noonan, Michael. 1985. Complementation. In Timothy Shopen (ed.), *Language typology and syntactic description. Complex constructions*, vol. 2, 42–140. Cambridge: Cambridge University Press.

Norcliffe, Elisabeth & Agnieszka E. Konopka. 2015. Vision and language in cross-linguistic research on sentence production. In Ramesh Kumar Mishra, Narayanan Srinivasan & Falk Huettig (eds.), *Attention and vision in language processing*. New York: Springer.

Osam, Emmanuel Kweku. 1994. Aspects of Akan grammar: a functional perspective. Oregon: University of Oregon PhD.

Papafragou, Anna. 2010. Source-goal asymmetries in motion representation: implications for language production and comprehension. *Cognitive Science* 34. 1064–1092. doi:10.1111/j.1551-6709.2010.01107.x.

Papafragou, Anna, Justin Hulbert & John Trueswell. 2008. Does language guide event perception? Evidence from eye movements. *Cognition* 108. 155–184.

Papafragou, Anna, Christine Massey & Lila Gleitman. 2002. Shake, rattle, "n" roll: the representation of motion in language and cognition. *Cognition* 84. 189–219.

Parril, Fey, Benjamin K. Bergen & Patricia V. Lichtenstein. 2013. Grammatical aspect, gesture, and conceptualization: Using co-speech gesture to reveal event representations. *Cognitive Linguistics* 24(1). 135–158.

Parsons, Terrence. 1990. *Events in the semantics of English: A study in subatomic semantics*. Cambridge, MA: MIT Press.

Pawley, Andrew. 1987. Encoding events in Kalam and English: different logics for reporting experience. In R. Tomlin (ed.), *Coherence and grounding in discourse*, vol. 11, 329–360. (Typological Studies in Language). Amsterdam/Philadelphia: John Benjamins.

Pawley, Andrew. 2008. Compact versus narrative serial verb constructions in Kalam. In Gunter Senft (ed.), *Serial verb constructions in Austronesian and Papuan languages*, 171–202. Canberra: Pacific Linguistics.

Pawley, Andrew. 2011. Event representation in serial verb constructions. In Jürgen Bohnemeyer & Eric Pederson (eds.), *Event representation in language and cognition*, 13–42. (Language Culture and Cognition 11). Cambridge: Cambridge University Press.

Pawley, Andrew & Johnathan Lane. 1998. From event sequence to grammar: Serial verb constructions in Kalam. In Anna Siewierska & Jae Jong Song (eds.), *Case, typology and grammar*, 201–227. Amsterdam/Philadelphia: John Benjamins.

Pine, Karen J., Nicola Lufkin & David Messer. 2004. More gestures than answers: Children learning about balance. *Developmental Psychology* 40(6). 1059–1067. doi:10.1037/0012-1649.40.6.1059.

Pinker, Steven. 1994. *The language instinct*. New York: William Morrow.

Pustejovsky, James. 1991. The syntax of event structure. In Beth Levin & Steven Pinker (eds.), *Lexical and conceptual semantics*, 47–82. Amsterdam: Elsevier Science Publishers.

Putten, Saskia van. 2009. Talking about motion in Avatime. Leiden University, Linguistics MPhil.

Putten, Saskia van. 2014a. Information structure in Avatime. Nijmegen: Radboud University Nijmegen PhD Thesis.

Putten, Saskia van. 2014b. Left-dislocation and subordination in Avatime (Kwa). In Rik Van Gijn, Jeremy Hammond, Dejan Matic, Saskia van Putten & Ana Vilacy Galucio (eds.), *Information structure and reference tracking in complex sentences*, 71–98. Amsterdam/Philadelphia: John Benjamins.

Quine, Willard Van Orman. 1985. Events and reification. In Ernest LePore & Brian P. McLaughlin (eds.), *Actions and events: perspectives on the philosophy of Donald Davidson*, 162–171. Oxford: Blackwell.

R Core Team. 2015. *R: A language and environment for statistical computing*. Vienna, Austria. http://www.R-project.org.

Radke, Richard J., Srinivas Andra, Omar Al-Kofahi & Badrinath Roysam. 2005. Image change detection algorithms: a systematic survey. *IEEE Transactions on Image Processing* 14(3). 294–307.

Regier, Terry. 1995. A model of the human capacity for categorizing spatial relations. *Cognitive Linguistics* 6. 63–88. doi:10.1515/cogl.1995.6.1.63.

Regier, Terry, Paul Kay & Naveen Khetarpal. 2007. Color naming reflects optimal partitions of color space. *Proceedings of the National Academy of Sciences* 104. 1436–1441. doi:10.1073/pnas.0610341104.

Regier, Terry & Mingyu Zheng. 2007. Attention to endpoints: A cross-linguistic constraint on spatial meaning. *Cognitive Science* 31. 705–719. doi:10.1080/15326900701399954.

Rifkin, Anthony. 1985. Evidence for a basic level in event taxonomies. *Memory & Cognition* 13(6). 538–556.

Roberson, Debi, Ian Davies & Jules Davidoff. 2000. Color categories are not universal: replications and new evidence from a stone-age culture. *Journal of Experimental Psychology: General* 129(3). 369–398.

Rosch, Eleanor, Carolyn B. Mervis, Wayne D. Gray, David M. Johnson & Penny Boyes-Braem. 1976. Basic objects in natural categories. *Cognitive Psychology* 8. 382–439.

Rosch, Eleanor, Carol Simpson & Scott R. Miller. 1976. Structural bases of typicality effects. *Journal of Experimental Psychology: Human Perception and Performance* 2(4). 491–502.

Rothstein, Susan. 2005. States and modification: A reply to Maienborn. *Theoretical Linguistics* 31(3). 375–381.

Ruiter, Jan Peter de. 2000. The production of gesture and speech. In David McNeill (ed.), *Language and gesture*, 248–311. Cambridge: Cambridge University Press.

Ryle, Gilbert. 1949. *The concept of mind*. London: Barnes and Noble.

Sargent, Jesse Q., Jeffrey M. Zacks, David Z. Hambrick, Rose T. Zacks, Christopher A. Kurby, Heather R. Bailey, Michelle L. Eisenberg & Taylor M. Beck. 2013. Event segmentation ability uniquely predicts event memory. *Cognition* 129. 241–255. doi:10.1016/j.cognition.2013.07.002.

Sasse, Hans-Jürgen. 2002. Recent activity in the theory of aspect: Accomplishments, achievements, or just non-progressive state? *Linguistic Typology* 6. 199–-271.

Schank, Roger C. & Robert P. Abelson. 1977. *Scripts, plans, goals and understanding: An inquiry into human knowledge structures*. Oxford: Lawrence Erlbaum.

Schuh, Russel G. 1995a. Aspects of Avatime phonology. *Studies in African Linguistics* 24(1). 31–67.

Schuh, Russel G. 1995b. Avatime noun classes and concord. *Studies in African Linguistics* 24(2). 123–149.

Schultze-Berndt, Eva. 2000. Simple and complex verbs in Jaminjung. A study of event categorisation in an Australian language. Nijmegen: Katholieke Universiteit Nijmegen PhD Dissertation.

Schwartz, Robert. 2008. Events are what we make of them. In Thomas F. Shipley & Jeffrey M. Zacks (eds.), *Understanding events: from perception to action*, 54–60. New York: Oxford University Press.

Sebba, Mark. 1987. *The syntax of serial verbs: an investigation into serialisation in Sranan and other languages*. Amsterdam/Philadelphia: John Benjamins.

Senft, Gunter. 2008. Event conceptualisation and event report in serial verb constructions in Kilivila: towards a new approach to research an old phenomenon. In Gunter Senft (ed.), *Serial verb constructions in Austronesian and Papuan languages*, 203–230. Canberra: Pacific Linguistics Publishers.

Seuren, Pieter. 1991. The definition of serial verbs. In Francis Byrne & Thom Huebner (eds.), *Development and structures of Creole languages: essays in honor of Derek Bickerton*, 193–205. Amsterdam/Philadelphia: John Benjamins.

Sharp, Richard & David I. Donaldson. 2007. Electrophysiological correlates of event segmentation: how does the human mind process ongoing activity? *Proceedings of the Annual Meeting of the Cognitive Neuroscience Society, New York*, 235. Cognitive Neuroscience Society.

Shipley, Thomas F. & Jeffrey M. Zacks (eds.). 2008. *Understanding Events*. (Oxford Series in Visual Cognition). Oxford: Oxford University Press.

Slobin, Dan I. 1987. Thinking for speaking. In J. Aske, N. Beery, L. Michaelis & H. Filip (eds.), *Proceedings of the 13th annual meeting of the Berkeley Linguistic Society meeting*, 435–445.

Slobin, Dan I. 1996a. From "thought and language" to "thinking for speaking." In John J. Gumperz & Stephen C. Levinson (eds.), *Rethinking linguistic relativity*, 70–96. (Studies in the Social and Cultural Foundations of Language 17). Cambridge: Cambridge University Press.

Slobin, Dan I. 1996b. Two ways to travel: Verbs of motion in English and Spanish. In M. Shibatani & S. A. Thompson (eds.), *Grammatical constructions: Their form and meaning*, 70–96. Oxford: Oxford University Press.

Slobin, Dan I. 2004. The many ways to search for a frog: Linguistic typology and the expression of motion events. In Sven Strömqvist & Ludo Verhoeven (eds.), *Relating events in narrative: Typological and contextual perspectives*, vol. 2, 219–257. Mahwah, NJ: Lawrence Erlbaum Associates.

Smith, Carlota S. 1999. Activities: states or events? *Linguistics and Philosophy* 22(5). 479–508.

Snodgrass, Joan Gay & Mary Vanderwart. 1980. A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory* 6(2). 174–215.

Staden, Miriam van & Ger Reesink. 2008. Serial verb constructions in a linguistic area. In Gunter Senft (ed.), *Serial verb constructions in Austronesian and Papuan languages*, 17–54. Canberra: Pacific Linguistics.

Stewart, John M. 1989. Kwa. In John Bendor-Samuel (ed.), *The Niger-Congo languages*, 217–246. Lanham: University Press of America.

Strickland, Brent & Frank C. Keil. 2011. Event completion: Event based inferences distort memory in a matter of seconds. *Cognition* 121. 409–415. doi:10.1016/j.cognition.2011.04.007.

Stutterheim, Christiane von & Ralf Nüse. 2003. Processes of conceptualisation in language production: language-specific perspectives and event construal. *Linguistics* 41(5). 851–881.

Swallow, Khena M., Deanna M. Barch, Denise Head, Corey J. Maley, Derek Holder & Jeffrey M. Zacks. 2011. Changes in events alter how people remember recent information. *Journal of Cognitive Neuroscience* 23(5). 1052–1064. doi:10.1162/jocn.2010.21524.

Swallow, Khena M., Jeffrey M. Zacks & Richard A. Abrams. 2009. Event boundaries in perception affect memory encoding and updating. *Journal of Experimental Psychology: General* 138(2). 236–257. doi:10.1037/a0015631.

Talmy, Leonard. 1985. Lexicalization Patterns: Semantic Structure in Lexical Forms. In Timothy Shopen (ed.), *Language Typology and Syntactic Description iii: Grammatical Categories and the Lexicon,* 36–149. Cambridge: Cambridge University Press.

Talmy, Leonard. 2000. *Towards a cognitive semantics*. . 2 vols. Cambridge, MA: MIT Press.

Trueswell, John & Anna Papafragou. 2010. Perceiving and remembering events cross-linguistically: Evidence from dual-task paradigms. *Journal of Memory and Language* 63. 64–82. doi:10.1016/j.jml.2010.02.006.

Tversky, Barbara. 1990. Where partonomies and taxonomies meet. In S. L. Tsohatzidis (ed.), *Meanings and prototypes: Studies in linguistic categorization*, 334–344. London: Routledge.

Van Valin, Robert. 2005. *Exploring the syntax-semantics interface*. Cambridge: Cambridge University Press.

Vendler, Zeno. 1957. Verbs and times. *Philosophical Review* 6. 143–60.

Verkerk, Annemarie. 2014. The evolutionary dynamics of motion event encoding. Radboud University Nijmegen Ph.D. Thesis.

Vigliocco, Gabriella, David P. Vinson, Federica Paganelli & Katharina Dworzynski. 2005. Grammatical gender effects on cognition: implications for language learning and language use. *Journal of Experimental Psychology: General* 134. 501–520.

Westermann, Diedrich. 1907. *Grammatik der Ewe-Sprache*. Berlin: Dietrich Reimer.

Wilder, David A. 1978a. Predictability of behaviours, goals, and unit of perception. *Personality and Social Psychology Bulletin* 4(4). 604–607.

Wilder, David A. 1978b. Effect of predictability on units of perception and attribution. *Personality and Social Psychology Bulletin* 4(2). 281–284.

Williamson, Kay & Roger Blench. 2000. Niger-Congo. In Bernd Heine & Derek Nurse (eds.), *African languages: an introduction*, 11–42. Cambridge: Cambridge University Press.

Winawer, Jonathan, Nathan Witthoft, Michael C. Frank, Lisa Wu, Alex R. Wade & Lera Boroditsky. 2007. Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Sciences of the USA* 104(19). 7780–7785.

Wolff, Phillip. 2003. Direct causation in the linguistic coding and indivuation of causal events. *Cognition* 88. 1–48.

Wolff, Phillip & Kevin J. Holmes. 2011. Linguistic relativity. *Cognitive Science* 2(3). 253–265.

Wolff, Phillip, Ga-Hyun Jeon & Yu Li. 2009. Causers in English, Korean, and Chinese and the individuation of events. *Language and Cognition* 1(2). 167–196.

Wolff, Phillip & Tatyana Ventura. 2009. When Russians learn English: How the semantics of causation may change. *Bilingualism: Language and Cognition* 12(2). 153–176.

Zacks, Jeffrey M. 2004. Using movement and intentions to understand simple events. *Cognitive Science* 28. 979–1008.

Zacks, Jeffrey M., Todd S. Braver, Margaret A. Sheridan, David I. Donaldson, Abraham Z. Snyder, John M. Ollinger, Randy L. Buckner & Marcus E. Raichle. 2001. Human brain activity time-locked to perceptual event boundaries. *Nature Neuroscience* 4. 651–655. doi:10.1038/88486.

Zacks, Jeffrey M., Nicole K. Speer, Khena M. Swallow, Todd S. Braver & Jeremy R. Reynolds. 2007. Event perception: A mind-brain perspective. *Psychological Bulletin* 133(2). 273–293. doi:10.1037/0033-2909.133.2.273.

Zacks, Jeffrey M., Nicole K. Speer, Jean M. Vettel & Larry L. Jacoby. 2006. Event understanding and memory in healthy aging and dementia of the Alzheimer type. *Psychology and Aging* 21(3). 466–482.

Zacks, Jeffrey M., Khena M. Swallow, Jean M. Vettel & Mark P. McAvoy. 2006. Visual motion and the neural correlates of event perception. *Brain Research* 1076. 150–162. doi:10.1016/j.brainres.2005.12.122.

Zacks, Jeffrey M. & Barbara Tversky. 2001. Event structure in perception and conception. *Psychological Bulletin* 127(1). 3–21. doi:10.1037/0033-2909.127.1.3.

Zacks, Jeffrey M., Barbara Tversky & Gowri Iyer. 2001. Perceiving, remembering and communicating structure in events. *Journal of Experimental Psychology: General* 130(1). 29–58. doi:10.1037//0096-3445.130.1.29.

## Samenvatting

Overal om ons heen vindt een continue stroom van activiteit plaats. Toch denken en praten we over gebeurtenissen alsof het discrete eenheden zijn. In dit proefschrift onderzoek ik de manieren waarop we in ons taalgebruik en ons denken gebeurtenissen tot eenheden maken en hoe deze eenheden met elkaar samenhangen. Hierbij focus ik op een specifieke syntactische constructie, de seriële werkwoordsconstructie (*serial verb construction,* afgekort SVC), in een specifieke taal, het Avatime. SVC's zijn constructies waarin twee of meer werkwoorden samengevoegd zijn in een enkel zinsdeel zonder dat er sprake is van subordinatie of coördinatie, zoals te zien in voorbeeld (1).

| | | | |
|---|---|---|---|
| *(1)* | *bị-lịla* | *kú* | *li-fu = nè* |
| | $C_4$p.PFV-verdwijn | binnengaan:LOC | $C_3$s-lucht = DEF |
| | 'Ze verdwenen de lucht in.' | | (Folkstory_110406_QM_01:48) |

Over deze constructies wordt vaak beweerd dat ze verwijzen naar enkelvoudige conceptuele gebeurtenissen. Er is echter in de literatuur weinig bewijs voor deze bewering. Dit proefschrift onderzoekt het verband tussen SVC's en de conceptualisatie van gebeurtenissen in het Avatime bij drie verschillende vormen van cognitie: denken voor spreken, geheugen en perceptie. Hierbij maak ik gebruik van verschillende methoden: beschrijvende taalkunde, gebarenanalyse en gedragsexperimenten.

In Hoofdstuk 1 bespreek ik eerder onderzoek naar het segmenteren van gebeurtenissen, het verband tussen taal en denken en seriële werkwoordsconstructies. Vervolgens formuleer ik de voornaamste onderzoeksvraag: Verwijzen Avatime SVC's naar datgene wat Avatime sprekers als enkelvoudige gebeurtenissen conceptualiseren? Ten slotte geef ik een overzicht van de drie gebieden van cognitie die in dit proefschrift onderzocht worden – denken voor spreken, perceptie en geheugen – en beschrijf ik de gebruikte onderzoeksmethoden.

Hoofdstuk 2 bevat een taalkundige beschrijving van het Avatime in het algemeen. Het vangt aan met een introductie van de sprekers, een groep van ongeveer 15 000 mensen in de Volta Regio van Ghana, en hun taal, een van de Ghana-Togo Mountain talen binnen de Kwa-tak van de Niger-Congo taalfamilie. Daarna volgt een overzicht van de grammatica van het Avatime. Van bijzonder

belang voor het onderzoek naar Avatime SVC's zijn de manieren waarop klinkercombinaties worden uitgesproken, de morfologie van werkwoorden en de verschillende soorten samengestelde zinnen.

Hoofdstuk 3 bevat een gedetailleerde beschrijving van de eigenschappen van Avatime SVC's. Het laat zien dat het Avatime drie soorten SVC's heeft. Net als in veel andere talen met seriële werkwoordsconstructies is er een onderscheid tussen constructies die naar opeenvolgende acties verwijzen en constructies die dit niet doen. Binnen de laatste groep is er ook een verschil tussen syntactisch meer geïntegreerde *nuclear-level* constructies die vooral voor modificatie worden gebruikt en minder geïntegreerde *core-level* constructies die gebruikt worden voor het manipuleren van de configuratie van argumenten. Dit laatste onderscheid is in de literatuur regelmatig gemaakt voor talen in de Austronesische regio, maar wordt bij West-Afrikaanse talen over het algemeen niet gemaakt.

In Hoofdstuk 4 onderzoek ik de relatie tussen SVC's en de conceptualisatie van gebeurtenissen tijdens het denken voor spreken. Dit onderzoek vergelijkt de timing van de gebaren die sprekers maken bij verschillende soorten constructies met meerdere werkwoorden. Het is gebaseerd op data uit vier monologen van in totaal 45 minuten. De resultaten laten zien dat SVC's vaak voorkomen met enkelvoudige gebaren die de volledige constructie overlappen en nooit met meerdere aparte gebaren. Andere constructies met meerdere werkwoorden komen daarentegen vaak voor met meerdere gebaren, waarbij voor elk werkwoord een apart gebaar wordt gebruikt. Een uitzondering hierop vormen complementszinnen met een infinitief werkwoord, die in dit opzicht meer op SVC's lijken. Deze bevindingen sluiten aan bij eerdere suggesties dat het zinsdeel de relevante eenheid is voor de formulering van de conceptuele boodschap. De resultaten suggereren dat sprekers van het Avatime enkelvoudige zinsdelen, waaronder SVC's, gebruiken om uitdrukking te geven aan datgene wat ze tijdens het denken voor spreken als enkelvoudige gebeurtenissen conceptualiseren.

In Hoofdstuk 5 onderzoek ik hoe gebeurtenissen gesegmenteerd worden op het moment dat ze worden waargenomen. De onderzoeksvragen zijn (i) of de gebeurtenissen die met SVC's worden beschreven op een andere manier gesegmenteerd worden dan andere gebeurtenissen en (ii) of het gebruik van SVC's de segmentatie tijdens de perceptie beïnvloedt. Aangezien SVC's worden gebruikt voor de beschrijving van opeenvolgende acties die een cultureel herkenbare eenheid

vormen, zullen mensen vaker SVC's gebruiken wanneer ze vertrouwde gebeurtenissen beschrijven dan wanneer ze onbekende gebeurtenissen beschrijven. In de literatuur is gesuggereerd dat vertrouwde gebeurtenissen grover gesegmenteerd worden tijdens de perceptie dan onbekende gebeurtenissen, maar de enige eerdere studie hiernaar vond geen invloed van vertrouwdheid. In twee experimenten test ik het effect van culturele vertrouwdheid op de segmentatie van gebeurtenissen bij sprekers van het Avatime en het Nederlands. De resultaten laten zien dat vertrouwdheid enige invloed heeft, maar dat deze afhangt van de lengte van de gebeurtenis. Langere gebeurtenissen werden grover gesegmenteerd wanneer ze vertrouwd waren dan wanneer ze onbekend waren. Korte gebeurtenissen werden echter juist fijner gesegmenteerd wanneer ze vertrouwd waren. Vertrouwdheid lijkt dus wel invloed te hebben op de segmentatie van gebeurtenissen tijdens de perceptie, maar hoe deze invloed precies werkt is onduidelijk en dient verder onderzocht te worden. Een derde experiment onderzocht of het gebruik van SVC's de segmentatie tijdens de perceptie kan beïnvloeden. Deelnemers werden geprimed met ofwel SVC's ofwel samengestelde (nevenschikkende) zinnen waarna ze een segmentatietaak uitvoerden. De verwachting was dat priming met SVC's zou leiden tot een meer holistische segmentatie. Dit was inderdaad het geval: deelnemers die met SVC's geprimed werden segmenteerden gebeurtenissen in grotere eenheden waarbij ze grotere delen van de actie in enkelvoudige concepten verenigden dan deelnemers die geprimed werden met samengestelde zinnen. Al met al lijkt het onwaarschijnlijk dat gebeurtenissen die vaak met SVC's beschreven worden, zoals bijvoorbeeld vertrouwde gebeurtenissen, inherent anders worden gesegmenteerd dan andere gebeurtenissen. Het gebruik van SVC's lijkt echter wel invloed te kunnen hebben op de manier waarop mensen gebeurtenissen tijdens het waarnemen segmenteren.

Hoofdstuk 6 beschrijft een serie geheugenexperimenten waarin onderzocht wordt hoe mensen gebeurtenissen onthouden die normaal gesproken met SVC's beschreven worden. Deze experimenten focussen op gebeurtenissen waarin objecten worden geplaatst en weggenomen. In het Avatime moet bij het beschrijven van plaatsing in sommige gevallen een SVC gebruikt worden, waarin het werkwoord voor plaatsing voorafgegaan wordt door een werkwoord dat wegnemen uitdrukt. Twintig plaatsingsgebeurtenissen die met een dergelijke SVC beschreven worden en twintig plaatsingsgebeurtenissen die niet met een SVC beschreven worden werden gefilmd. Voor elke plaatsingsgebeurtenis werd ook een voorafgaande

wegneemgebeurtenis gefilmd. Sprekers van het Avatime, het Nederlands en het Engels kregen van elk wegneem-plaats-paar ofwel het wegnemen ofwel het plaatsen te zien. Daarna bekeken ze alle video's en gaven ze aan of ze de video eerder hadden gezien of niet. De verwachting was dat sprekers van het Avatime de wegneem-plaats episodes die met SVC's beschreven worden als enkelvoudige eenheden zouden onthouden en daarom bij deze episodes vaker onterecht een niet eerder bekeken video als bekeken zouden herkennen. De resultaten laten echter zien dat de gebeurtenissen die in het Avatime niet met SVC's beschreven worden door sprekers van zowel het Engels, het Nederlands als het Avatime beter onthouden worden. Dit resultaat suggereert een algemeen verschil in hoe mensen gebeurtenissen onthouden dat in het Avatime in SVC's gereflecteerd wordt.

In hoofdstuk 7 vat ik de belangrijkste bevindingen van dit proefschrift samen en kom ik terug op de belangrijkste thema's, namelijk of SVC's naar enkelvoudige gebeurtenissen verwijzen en wat dit betekent en wat de relatie is tussen de segmentatie van gebeurtenissen in taal en cognitie. SVC's en enkelvoudige zinnen in het algemeen blijken inderdaad naar enkelvoudige conceptuele gebeurtenissen te verwijzen tijdens het denken voor spreken. Ook tijdens de perceptie is er invloed van SVC's, wat blijkt uit het resultaat dat priming met SVC's tot een meer holistisch segmentatiepatroon leidde. De verwachting was dat er als gevolg van deze invloed tijdens de perceptie ook een invloed zou zijn op het geheugen voor gebeurtenissen, maar hiervoor werd geen bewijs gevonden. Al met al laat dit proefschrift een connectie zien tussen SVC's en een meer holistische conceptualisatie van gebeurtenissen, alsmede een algemene connectie tussen de segmentatie van gebeurtenissen in taal en in cognitie. Deze connectie is het sterkst tijdens het spreken, maar is ook aanwezig tijdens het waarnemen van gebeurtenissen.

## Biographical note

Rebecca Defina studied a combined BSc in advanced mathematics and BA majoring in linguistics at the University of Sydney and graduated in 2007. She then went on to do an MPhil in linguistics at Leiden University. During her MPhil she received a Field Trip Grant from the Endangered Languages Documentation Programme for the description and documentation of Avatime. Her thesis describing the tense, aspect, and mood system of Avatime was completed in 2009 and she graduated cum laude. In 2010 she began her PhD at the Max Planck Institute for Psycholinguistics in the Language and Cognition group. There she continued her work on Avatime, expanding the grammatical description and pursuing an in-depth investigation of serial verb constructions and their connection to conceptual events. This led her to develop a more experimental focus and also to conduct comparative studies with Dutch and English speakers. She currently works as a postdoctoral researcher with the Centre of Excellence for the Dynamics of Language at the University of Melbourne where she is studying how Pitjantjatjara speaking children learn serial verb constructions and the events they describe. Her research interests include syntax, semantics, and the relations between language, culture, and cognition.

## MPI Series in Psycholinguistics

1. The electrophysiology of speaking: Investigations on the time course of semantic, syntactic, and phonological processing. *Miranda van Turennout*
2. The role of the syllable in speech production: Evidence from lexical statistics, metalinguistics, masked priming, and electromagnetic midsagittal articulography. *Niels O. Schiller*
3. Lexical access in the production of ellipsis and pronouns. *Bernadette M. Schmitt*
4. The open-/closed-class distinction in spoken-word recognition. *Alette Haveman*
5. The acquisition of phonetic categories in young infants: A self-organising artificial neural network approach. *Kay Behnke*
6. Gesture and speech production. *Jan-Peter de Ruiter*
7. Comparative intonational phonology: English and German. *Esther Grabe*
8. Finiteness in adult and child German. *Ingeborg Lasser*
9. Language input for word discovery. *Joost van de Weijer*
10. Inherent complement verbs revisited: Towards an understanding of argument structure in Ewe. *James Essegbey*
11. Producing past and plural inflections. *Dirk Janssen*
12. Valence and transitivity in Saliba: An Oceanic language of Papua New Guinea. *Anna Margetts*
13. From speech to words. *Arie van der Lugt*
14. Simple and complex verbs in Jaminjung: A study of event categorisation in an Australian language. *Eva Schultze-Berndt*
15. Interpreting indefinites: An experimental study of children's language comprehension. *Irene Krämer*
16. Language-specific listening: The case of phonetic sequences. *Andrea Weber*
17. Moving eyes and naming objects. *Femke van der Meulen*
18. Analogy in morphology: The selection of linking elements in Dutch compounds. *Andrea Krott*
19. Morphology in speech comprehension. *Kerstin Mauth*

20. Morphological families in the mental lexicon. *Nivja H. de Jong*
21. Fixed expressions and the production of idioms. *Simone A. Sprenger*
22. The grammatical coding of postural semantics in Goemai (a West Chadic language of Nigeria). *Birgit Hellwig*
23. Paradigmatic structures in morphological processing: Computational and cross-linguistic experimental studies. *Fermín Moscoso del Prado Martín*
24. Contextual influences on spoken-word processing: An electrophysiological approach. *Daniëlle van den Brink*
25. Perceptual relevance of prevoicing in Dutch. *Petra M. van Alphen*
26. Syllables in speech production: Effects of syllable preparation and syllable frequency. *Joana Cholin*
27. Producing complex spoken numerals for time and space. *Marjolein Meeuwissen*
28. Morphology in auditory lexical processing: Sensitivity to fine phonetic detail and insensitivity to suffix reduction. *Rachèl J. J. K. Kemps*
29. At the same time...: The expression of simultaneity in learner varieties. *Barbara Schmiedtová*
30. A grammar of Jalonke argument structure. *Friederike Lüpke*
31. Agrammatic comprehension: An electrophysiological approach. *Marlies Wassenaar*
32. The structure and use of shape-based noun classes in Miraña (North West Amazon). *Frank Seifart*
33. Prosodically-conditioned detail in the recognition of spoken words. *Anne Pier Salverda*
34. Phonetic and lexical processing in a second language. *Mirjam Broersma*
35. Retrieving semantic and syntactic word properties. *Oliver Müller*
36. Lexically-guided perceptual learning in speech processing. *Frank Eisner*
37. Sensitivity to detailed acoustic information in word recognition. *Keren B. Shatzman*
38. The relationship between spoken word production and comprehension. *Rebecca Özdemir*
39. Disfluency: Interrupting speech and gesture. *Mandana Seyfeddinipur*

40. The acquisition of phonological structure: Distinguishing contrastive from non-contrastive variation. *Christiane Dietrich*
41. Cognitive cladistics and the relativity of spatial cognition. *Daniel B.M. Haun*
42. The acquisition of auditory categories. *Martijn Goudbeek*
43. Affix reduction in spoken Dutch. *Mark Pluymaekers*
44. Continuous-speech segmentation at the beginning of language acquisition: Electrophysiological evidence. *Valesca Kooijman*
45. Space and iconicity in German Sign Language (DGS). *Pamela Perniss*
46. On the production of morphologically complex words with special attention to effects of frequency. *Heidrun Bien*
47. Crosslinguistic influence in first and second languages: Convergence in speech and gesture. *Amanda Brown*
48. The acquisition of verb compounding in Mandarin Chinese. *Jidong Chen*
49. Phoneme inventories and patterns of speech sound perception. *Anita Wagner*
50. Lexical processing of morphologically complex words: An information-theoretical perspective. *Victor Kuperman*
51. A grammar of Savosavo, a Papuan language of the Solomon Islands. *Claudia Wegener*
52. Prosodic structure in speech production and perception. *Claudia Kuzla*
53. The acquisition of finiteness by Turkish learners of German and Turkish learners of French: Investigating knowledge of forms and functions in production and comprehension. *Sarah Schimke*
54. Studies on intonation and information structure in child and adult German. *Laura de Ruiter*
55. Processing the fine temporal structure of spoken words. *Eva Reinisch*
56. Semantics and (ir)regular inflection in morphological processing. *Wieke Tabak*
57. Processing strongly reduced forms in casual speech. *Susanne Brouwer*
58. Ambiguous pronoun resolution in L1 and L2 German and Dutch. *Miriam Ellert*
59. Lexical interactions in non-native speech comprehension: Evidence from electro-encephalography, eye-tracking, and functional magnetic resonance imaging. *Ian FitzPatrick*

60. Processing casual speech in native and non-native language. *Annelie Tuinman*
61. Split intransitivity in Rotokas, a Papuan language of Bougainville. *Stuart Robinson*
62. Evidentiality and intersubjectivity in Yurakaré: An interactional account. *Sonja Gipper*
63. The influence of information structure on language comprehension: A neurocognitive perspective. *Lin Wang*
64. The meaning and use of ideophones in Siwu. *Mark Dingemanse*
65. The role of acoustic detail and context in the comprehension of reduced pronunciation variants. *Marco van de Ven*
66. Speech reduction in spontaneous French and Spanish. *Francisco Torreira*
67. The relevance of early word recognition: Insights from the infant brain. *Caroline Junge*
68. Adjusting to different speakers: Extrinsic normalization in vowel perception. *Matthias J. Sjerps*
69. Structuring language. Contributions to the neurocognition of syntax. *Katrien R. Segaert*
70. Infants' appreciation of others' mental states in prelinguistic communication: A second person approach to mindreading. *Birgit Knudsen*
71. Gaze behavior in face-to-face interaction. *Federico Rossano*
72. Sign-spatiality in Kata Kolok: how a village sign language of Bali inscribes its signing space. *Conny de Vos*
73. Who is talking? Behavioural and neural evidence for norm-based coding in voice identity learning. *Attila Andics*
74. Lexical processing of foreign-accented speech: Rapid and flexible adaptation. *Marijt Witteman*
75. The use of deictic versus representational gestures in infancy. *Daniel Puccini*
76. Territories of knowledge in Japanese conversation. *Kaoru Hayano*
77. Family and neighbourhood relations in the mental lexicon: A cross-language perspective. *Kimberley Mulder*
78. Contributions of executive control to individual differences in word production. *Zeshu Shao*

79. Hearing speech and seeing speech: Perceptual adjustments in auditory-visual processing. *Patrick van der Zande*
80. High pitches and thick voices: The role of language in space-pitch associations. *Sarah Dolscheid*
81. Seeing what's next: Processing and anticipating language referring to objects. *Joost Rommers*
82. Mental representation and processing of reduced words in casual speech. *Iris Hanique*
83. The many ways listeners adapt to reductions in casual speech. *Katja Poellmann*
84. Contrasting opposite polarity in Germanic and Romance languages: Verum Focus and affirmative particles in native speakers and advanced L2 learners. *Giuseppina Turco*
85. Morphological processing in younger and older people: Evidence for flexible dual-route access. *Jana Reifegerste*
86. Semantic and syntactic constraints on the production of subject-verb agreement. *Alma Veenstra*
87. The acquisition of morphophonological alternations across languages. *Helen Buckler*
88. The evolutionary dynamics of motion event encoding. *Annemarie Verkerk*
89. Rediscovering a forgotten language. *Jiyoun Choi*
90. The road to native listening: Language-general perception, language-specific input. *Sho Tsuji*
91. Infants'understanding of communication as participants and observers. *Gudmundur Bjarki Thorgrímsson*
92. Information structure in Avatime. *Saskia van Putten*
93. Switch reference in Whitesands. *Jeremy Hammond*
94. Machine learning for gesture recognition from videos. *Binyam Gebrekidan Gebre*
95. Acquisition of spatial language by signing and speaking children: a comparison of Turkish sign language (TID) and Turkish. *Beyza Sümer*
96. An ear for pitch: on the effects of experience and aptitude in processing pitch in language and music. *Salomi Savvatia Asaridou*

97. lncrementality and Flexibility in Sentence Production. *Maartje van de Velde*
98. Social learning dynamics in chimpanzees: Reflections on (nonhuman) animal culture. *Edwin van Leeuwen*
99. The request system in Italian interaction. *Giovanni Rossi*
100. Timing turns in conversation: A temporal preparation account. *Lilla Magyari*
101. Assessing birth language memory in young adoptees. *Wencui Zhou*
102. A social and neurobiological approach to pointing in speech and gesture. *David Peeters*
103. Investigating the genetic basis of reading and language skills. *Alessandro Gialluisi*
104. Conversation Electrified: The Electrophysiology of Spoken Speech Act Recognition. *Rósa Signý Gisladottir*
105. Modelling Multimodal Language Processing. *Alastair Smith*
106. Predicting language in different contexts: The nature and limits of mechanisms in anticipatory language processing. *Florian Hintz*
107. Situational variation in non-native communication. *Huib Kouwenhoven*
108. Sustained attention in language production. *Suzanne Jongman*
109. Acoustic reduction in spoken-word processing: Distributional, syntactic, morphosyntatic, and orthographic effects. *Malte Viebahn*
110. Nativeness, dominance, and the flexibility of listening to spoken language. *Laurence Bruggeman*
111. Semantic specificity of perception verbs in Maniq. *Ewelina Wnuk*
112. On the identification of FOXP2 gene enhancers and their role in brain development. *Martin Becker*
113. Events in language and thought: The case of serial verb constructions in Avatime. *Rebecca Defina*