

©Copyright 2019

Joshua Crowgey

Braiding Language (by Computer): Lushootseed Grammar Engineering

Joshua Crowgey

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2019

Reading Committee:

Emily M. Bender, Chair

David Beck

Sharon Hargus

Program Authorized to Offer Degree:
Linguistics

University of Washington

Abstract

Braiding Language (by Computer): Lushootseed Grammar Engineering

Joshua Crowgey

Chair of the Supervisory Committee:
Professor Emily M. Bender
Linguistics

This dissertation describes the beginnings of the *təbšucid* project. *təbšucid*, literally the braiding of language, is a way to refer to “grammar” in Lushootseed (also known as Puget Salish, ISO-639-3:lut). The *təbšucid* project has three overlapping goals: (1) to advance linguistic science via grammar engineering methods with a specific focus on Lushootseed; (2) to package and distribute linguistic results in a way which is useful for people involved in the development of language-related applications which may serve a role in the documentation and revitalization of endangered languages; (3) to highlight the inherent value of the traditional language of the Puget Sound. To those ends, I began to build a system which could process Lushootseed texts. I implemented an initial morphophonological analyzer which can map Lushootseed orthography to a regularized morphophonemic representation. This representation is one which can serve as input to a syntactico-semantic grammar which provides semantic analyses for input sentences. I then implemented an initial syntactico-semantic grammar which maps the morpheme-regularized form to sentence-level semantics, via an explicit syntactic representation. In doing so, I address the first and third motivations listed above by presenting a series of case studies which emerged from the initial implementation work. In connection to the second motivation, this dissertation describes both the overall architecture and the context of the system, in the hope that the work I’ve carried out to-date can be something which can be built upon by others.

TABLE OF CONTENTS

	Page
List of Figures	iv
List of Tables	v
Chapter 1: Introduction	1
1.1 Motivation	2
1.2 About the author	7
1.3 Introduction to the Dissertation	9
Chapter 2: Lushootseed	15
2.1 Language Name and Areal Distribution	15
2.2 History	19
2.3 Philogenetics	36
2.4 Typology	39
2.5 Conclusion	47
Chapter 3: Linguistic Theory and Framework	48
3.1 Bird’s Eye View	48
3.2 Finite State Morphology and Phonology	50
3.3 Head-driven Phrase Structure Grammar for Morphosyntax	58
3.4 Introduction to DELPH-IN and Type Description Language	69
3.5 Minimal Recursion Semantics	71
3.6 Conclusion	85
Chapter 4: Environment and Development Cycle	86
4.1 Overall Architecture	86
4.2 The LinGO Grammar Matrix	88

4.3	Metagrammar Engineering: the CLIMB approach	91
4.4	Meta-choices	92
4.5	Transducer Generation	95
4.6	Testsuites, Grammar profiles, Regression Testing	98
4.7	Lushootseed data, testsuite construction	100
4.8	Conclusion	102
Chapter 5:	Implemented Lushootseed Morphophonology	104
5.1	Resources, Methodology, Initial Results	105
5.2	Structures	113
5.3	Relationship to Descriptive/Documentary Efforts	123
5.4	Conclusion	125
Chapter 6:	Grammar Engineering Examples	126
6.1	VSOv2 Word Order	126
6.2	A morphological example: object markers	149
6.3	Conclusion	157
Chapter 7:	Case-study I: Valence-increasing Morphology	158
7.1	Data	159
7.2	Theoretical and Practical Background	175
7.3	Analysis	201
7.4	Results	228
7.5	Conclusion	232
Chapter 8:	Case Study II: Verbal Arguments and Non-verbal Predicates	236
8.1	Nouns and Verbs in Salishan	237
8.2	Data	241
8.3	Analysis	251
8.4	Discussion and Results on the Testsuite	278
Chapter 9:	Evaluation	280
9.1	Evaluation Context: other notable implementations	281
9.2	Results	298

9.3 Error Analysis on Unparsed Items	316
9.4 Conclusion	328
Chapter 10: Conclusion	329
Bibliography	331
Appendix A: Scripts	352
Appendix B: Results for Valence Increasing Morphology Testsuite	356

LIST OF FIGURES

Figure Number	Page
2.1 Lushootseed speaking area	17
3.1 Automaton to model the non-finite whimsical language of bovines <i>/moo*</i> /.	52
3.2 Transducer mapping the whimsical language of bovines <i>/moo*</i> / to a whimsical language of specters <i>/boo*</i> /.	53
3.3 Transducers implementing Phonological Transformation Rules	56
3.4 Automata to model the language of three temporal proclitics	56
3.5 Automaton created by concatenating three regular languages	56
3.6 Transducer created by composition of the automaton in Figure 3.5 with the transducer of Figure 3.3.	57
3.7 A hierarchy of animals	60
3.8 An animals hierarchy with feature inheritance and typed feature values.	61
4.1 Text-processsing system architecture	87
4.2 The text processing system as a pipeline.	88
4.3 The customization of a grammar	90
5.1 Enumeration of the FST development methodology	108
5.2 Initial Performance of Morphophonological Analyzer	112
5.3 Network overview for the morphotactic system	119
8.1 A set of tense, aspect and mood choices implementing Lushootseed general prefixes.	248

LIST OF TABLES

Table Number	Page
1.1 Abbreviations for linguistic glosses used in this document	14
2.1 Selected Lushootseed place names and their English corollaries. Mostly from (Hilbert et al., 2001)	18
2.2 Salishan Genealogical Grouping by Swadesh (1950)	37
2.3 Salishan Genealogical Grouping by Thompson (1979)	38
2.4 Lushootseed Phonological Inventory	40
2.5 Lexical Suffixes used as numeral classifiers (Hess, 1995, 20)	43
2.6 Referents of Lushootseed lexical suffixes: the same referent may require different suffixes in different contexts (Hess, 1998, 22–23).	46
2.7 Lushootseed Nouns with corresponding lexical suffixes	47
3.1 The Chomsky-Schützenberger hierarchy	50
3.2 Structured Labels in DELPH-IN Elementary Predications	73
4.1 [incr tsdb()] fields and the corresponding data types available for each test- suite item	99
4.2 Testsuites used in HPSG Grammar Development	103
5.1 The FST evaluation data to date	112
5.2 Affixal template of Lushootseed verb stems	114
6.1 The forms of Lushootseed main clause subject markers	131
6.2 Testsuite items for subject markers and general word order constraints	132
6.3 Results against the extended testsuite	148
6.4 Lushootseed object markers	149
6.5 Development testsuite for Lushootseed Object Markers	151
7.1 Properties contributing to AGENT and PATIENT proto-roles	186
7.2 The dimensions of Hopper and Thompson’s (1980) transitivity scale	187
7.3 Summary of grammar coverage over valence-increasing morphology testsuite	229

8.1	Lushootseed General Prefixes	246
8.2	The Testsuite Results for System I	279
8.3	The Testsuite Results for System II	279
9.1	The evaluation results for Bear and Ant	301
9.2	The evaluation results for Bear and Fishhawk	307
9.3	The evaluation results for Coyote and Big Rock	313
9.4	The evaluation results for Mink	315

ACKNOWLEDGMENTS

I would first like to humbly acknowledge and thank the Lushootseed-speaking elders who shared their language and recorded their stories. Although I never knew them personally, the voices of Vi *taqʷšəblu* Hilbert, Edward *sʔadacut* Sam, Martha *səswixab* Lamont and others have changed me and influenced who I am and how I think about the world.

I would like to thank Emily M. Bender, a language scientist *par excellence*—indefatigable in her search for generalizations and explanations while never taking her eyes far from the rich diversity of the data she examines. She carves mystifying phenomena into empirical questions, seemingly with ease.

This thesis would not have been possible without massive amounts of help and guidance from David Beck. His deep knowledge of Lushootseed grammatical structures often steered me away from analytical dead ends and connundra.

I would also like to thank Sharon Hargus, whose careful reading improved this dissertation enormously.

In addition to the elder Lushootseed speakers who shared their language, in some cases, before I was born, I owe a huge debt of gratitude to the contemporary Lushootseed community who shared with me their classes, their dinner tables and their friendship. My heart overflows with gratitude to Angela *yuʔyuʔbəč stadəy?* Wymer, who opened her Lushootseed classroom to me in 2013, inviting me to learn not only Lushootseed words and grammar, but also Lushootseed culture and values.

I owe a similar debt of gratitude to Zalmai *ʔəswəli* Zahir, who not only accepted me into his online weekly classes for language learners. *ʔəswəli* continues to refine my thinking about *təbšucid*: Lushootseed grammar, the braiding of language.

I also would like to thank Jill LaPointe, Vi Hilbert's granddaughter and the President of the board of Lushootseed Research. She met with me in 2013 when I was just beginning my Lushootseed journey. When I asked her about how to access knowledge and materials, she invited me to the annual Lushootseed Language conference where I met Jay Miller, Crisca Bierwort, Pamela Amoss and other scholars whose work I ended up relying on.

Laurel Sercombe met with me and discussed the collection of Lushootseed recordings in the University of Washington Ethnomusicology library. John Vallier also helped me understand the nature of Arthur Ballard's legacy, providing insight into library records which I had previously misunderstood.

Additionally, I would like to acknowledge guidance from Stacy Waters and Walter Andrews. These two scholars not only provided me with interesting work on the Newbook Project in Digital Humanities, their depth of experience at the University of Washington kept my academic trajectory on the path to success.

During the time in which I was carrying out of this research, I received support from the Department of Linguistics. I would like to thank the Department for their official support of my work. Within the Department of Linguistics, Joyce Parvi was the first person I ever spoke to. I approached the office out of the blue in 2007 with a few questions about how to take a class. It was in part due to Joyce Parvi's encouragement to sign up for non-matriculated classes that got my feet wet in Linguistic studies and changed the course of my life. I'd also like to thank Michael Furr, who not only encouraged me to stay the course, but did so in a way which left me happily chuckling. Thank you, Mike. In the course of my studies in the Department, I benefited from the tutelage of world-renowned teachers who challenged me. I would like to thank each of my teachers, especially the ones I disagreed with.

I was also supported during the time of this dissertation by the Language Learning Center. Thank you to Paul Aoki for hiring me. Thank you to Russel Hugo for advocating for me. Thank you to Martin Eller for teaching me how to use strcopy.

This dissertation would never have been completed if it hadn't been for the work of Lisa Harper, my partner and wife, who took on so many of my responsibilities while I was whittling away in the corner on a really, really long paper. Thank you, Lisa. Can we share the credential?

Finally, to Bird, my friend of 11 years who passed away recently: thank you for showing me so many times over the years what love really means.

DEDICATION

to Lisa Harper, my north star

Chapter 1

INTRODUCTION

This dissertation describes the beginnings of the *təbšucid* project. *təbšucid*, literally the braiding of language, is a way to refer to “grammar” in Lushootseed (also known as Puget Salish, ISO-639-3:lut). The *təbšucid* project has three overlapping goals: (1) to advance linguistic science via grammar engineering methods with a specific focus on Lushootseed; (2) to package and distribute linguistic results in a way which is useful for people involved in the development of language-related applications which may serve a role in the documentation and revitalization of endangered languages; (3) to highlight the inherent value of the traditional language of the Puget Sound region.¹ To those ends, I began to build a system which could process Lushootseed texts. I implemented an initial morphophonological analyzer which can map Lushootseed orthography to a regularized morphophonemic representation. This representation is one which can serve as input to a syntactico-semantic grammar which provides semantic analyses for input sentences. I then implemented an initial syntactico-semantic grammar which maps the morpheme-regularized form to sentence-level semantics, via an explicit syntactic representation. In doing so, I address the first and third motivations listed above by presenting a series of case studies which emerged from the initial implementation work. In connection to the second motivation, this dissertation describes both the overall architecture and the context of the system, in the hope that the work I’ve carried out to-date is something which can be built upon by others.

The rest of this chapter comprises three sections. In the first, I dig a bit more into each of the motivations above in order to more precisely ground this work in the context I view it in.

¹Puget Sound is the current name for a relatively large protected saltwater body on the Pacific Northwest coast of North America. Chapter 2 provides further introduction to the area where Lushootseed is spoken.

Following that, I introduce the contents of this in more detail, laying out the topics addressed and the overall takeaways. Within that latter section are other utilitarian details such as the list of linguistic glosses, some conventions used in references etc.

1.1 Motivation

The three motivations presented at the beginning of this chapter might be restated as (1) Implemented grammars are good for Linguistics, (2) Linguistics can be good for others, (3) Linguistic Diversity is good for Linguistic Science. This section contains a short essay on each of these points which is intended to link my work securely to the context and content which inspired it.

1.1.1 Implemented Grammars are Good for Linguistics

The field of Grammar Engineering seeks to encode grammatical descriptions into software. This line of research can be seen as a multi-decade effort towards putting into practice one of the foundational principles of generative grammar, as expressed in the earlier writings of Chomsky (*inter alia*). For example, Chomsky (1965, 4) focuses on the need for explicitness and rigor.

A grammar of a language purports to be a description of the ideal speaker-hearer's intrinsic competence. If the grammar is, furthermore, perfectly explicit — in other words, if it does not rely on the intelligence of the understanding reader but rather provides an explicit analysis of his contribution — we may (somewhat redundantly) call it a generative grammar.

Later in the same work, Chomsky emphasizes that 20th century results in the foundations of mathematics have provided the formal techniques for characterizing language's recursive properties, opening doors through which contemporary research could proceed where earlier linguists could only dimly speculate:

But the fundamental reason of this inadequacy of traditional grammars is a more technical one. Although it was well understood that linguistic processes are in some

sense “creative,” the technical devices for expressing a system of recursive processes were simply not available until much more recently. In fact, a real understanding of how a language can (in Humboldt’s words) “make infinite use of finite means” has developed only within the last thirty years, in the course of studies in the foundations of mathematics. Now that these insights are readily available it is possible to return to the problems that were raised, but not solved, in traditional linguistic theory, and to attempt an explicit formulation of the “creative” processes of languages. There is, in short, no longer a technical barrier to the full-scale study of generative grammars. (Chomsky, 1965, 8)

While it has been said (Sag, 2010, *inter alia*) that in later years Chomsky’s own work began to abandon the explicitness criterion set out in earlier work, many linguists have continued in this spirit, both in focusing on the preciseness of description and in the co-opting of emerging technology to push the boundaries of what can be accounted for in linguistic theory. Fast-forwarding through much of this work to the present day, contemporary linguists are using computational infrastructure to build integrated systems of morphophonology (Beesley and Karttunen, 2003, *inter alia*) as well as morphosyntax and semantics (Flickinger, 2000, *inter alia*). Bender and her coauthors (Bender and Langendoen, 2010; Bender and Good, 2010) describe how computational techniques allow for more-efficient and scalable linguistic work.² These authors point out that much of the groundwork has already been laid for new, integrated models. Linguistic analysis, when implemented in a computer environment, forces a down-to-the-bits encoding of the analysis — essentially ruling out the vagueness of “relying on the intelligence of the understanding reader” which Chomsky bemoans in the traditional grammatical descriptions he references.

In Bender 2008, arguments for the merits of implemented grammars as linguistic hypothesis testers focus on the real-world complexity of natural language syntactic structures when all phenomena interacting in a sentence have to be accounted for. In Bender et al. 2011, the

²In Bender and Langendoen (2010), the authors connect the efficiency claims to the pressing issue of linguistic extinction (Hale, 1992; Krauss, 1992), an extremely salient point in the context of Lushootseed, a highly endangered language.

authors present a practical overview of the workflow of a grammar engineer as part of a discovery methodology for linguistic structures, describing the tools of the trade and the gains in scope and efficiency that they produce. Bender 2010 describes a case study which in many ways provides a model for the current project, implementing a machine-readable grammar for Wambaya, a non-Pama-Nyungan language of Australia (ISO:639-3:wmb), based on description and documentation from fieldwork. This latter work also highlights the fecundity of collaborations between linguists trained as grammar engineers and fieldworkers who collect primary data from speakers. Amongst the merits of the methodology presented in these works are the rapid identification of new phenomena across large corpora as well as the rigorous testing of interactions between subsystems of the grammar.

Connected to these arguments for using implemented grammars to improve the quality and efficiency of linguistic analysis are papers analyzing the reproducibility of linguistic analysis. Both Fokkens et al. 2013 and Maxwell 2012 emphasize the reproducibility of results as a cornerstone of science, the latter with a specific focus on descriptive grammars as research artifacts. Maxwell argues that the descriptive properties of linguistic analyses should be validated against a test suite of real data and crucially, that the data be made available. Fokkens et al. (2013) provide a case-study in reproducing results in Natural Language Processing (NLP) papers. Although the work of these authors highlights that using an implemented grammar is not the only required step to ensure reproducibility, it is a necessary one within the context of grammatical analysis.

1.1.2 Linguistics can be Good for Others

On the second motivation, the packaging and distribution of the results of linguistic analysis so that they can be used by others, I note that systems which encode linguistic knowledge form the backbone of many NLP tools, many of which have become common for resource-rich, economically powerful languages. These tools include spell-checkers, grammar-checkers and other programs intended to aid in writing, as well as programs meant to query the world-wide web for information or to provide machine translation services. In recent years, speech-to-text systems provide a key component of voice-driven human-computer interfaces, but such systems are not

typical for endangered languages.

However, just as the end-user applications for NLP listed above often rely on a backbone of implemented linguistic analysis, applied linguistic analysis itself relies on the availability of data regarding the language(s) in question: descriptions, recordings, interlinear glossed text (IGT), i.e. the products of fieldwork and traditional analysis. Computational linguists have identified a resource-poor/resource-rich dichotomy in which minority and lesser-spoken languages often do not have the dictionaries, descriptions and recordings necessary to bootstrap the development of an implemented linguistic system, much less the treebanks and electronic lexica which are widely available for English, French, German or Japanese. Some linguists have developed innovative algorithms and procedures for squeezing as much information as possible about a language out of minimal resources (Xia and Lewis, 2009; Bender et al., 2013).³ However, in the ideal case, linguists working on a given language aim to implement and distribute their work in a way that can be exported, downloaded and utilized by others.⁴ A focus on usability not only falls in line with scientific ideals, it also magnifies the impact of research on real human lives in the same way that innovations in chemistry and physics can lead to the eventual development of life-saving medical procedures. I suggest that the proper export and packaging of linguistic research may form the first steps in developing technologies that aid in the conservation and revitalization of languages which are on the brink of extinction.

1.1.3 Linguistic Diversity is Good for Linguistic Science

The third motivation of this project is to emphasize the value of the indigenous languages of the Americas and to focus linguistic study on one understudied, disappearing language of Western Washington.

In surveying the state-of-affairs with respect to the world's linguistic diversity, Hale (1992) provides a warning that the loss of languages is also related to a larger process of "loss of cultural

³<http://depts.washington.edu/uwcl/aggregation/>

⁴See Bird and Simons (2003).

and intellectual diversity in which politically dominant languages and cultures simply overwhelm indigenous local languages and cultures.” Furthermore, attitudes and beliefs about the value of indigenous language and cultures feeds into the process of language death (Hugo, 2010). I believe that committing linguistic efforts to the study of indigenous languages and cultures goes against this disheartening trend by emphasizing their valuable and important contribution in the study of human culture and behavior.

I can go further, I think, and say that linguistic science relies on the existence of linguistic diversity. The structures of the human languages of the world are the data which the field of linguistics seeks to describe. Yet, as Hale (1992) warns, the rate of language loss in contemporary times is unparalleled in history, estimating that roughly 50% of the world’s languages may fall silent within the 21st century, so efforts to document the current diversity of linguistic structures must be timely in order to capture as much information as possible about still-extant languages which may soon fall silent. Because of the efficiency and robustness of implemented linguistic analysis as described above, computational methodology can also form a valuable aid in the process of documenting and conserving endangered languages. This topic formed the basis of a workshop which began at the Association for Computational Linguistics meeting of 2014.⁵

1.1.4 Lushootseed study at the University of Washington

Linguistics diversity is a result of the speciation of tongues in connection to particular places on the earth as people have migrated and settled and lost intercommunication with erstwhile neighbors. Thus, linguistic diversity is connected strongly with place. As each place on the globe has a unique history, in many cases, there is a unique language and culture as well. This dissertation project was carried out under the auspices of the Department of Linguistics at the University of Washington in Seattle, an institution which resides on the place where $dx^w d\acute{a}w?abš$ (“Duwamish”) and $š\acute{a}č\acute{u}?abš$ (“Lake”) people spoke their language for many centuries: their language is Lushootseed. Because of this connection, I believe that studying Lushootseed at the

⁵<https://computel-workshop.org/>

University of Washington in Seattle is an important way to honor the specific place where the University resides.

The University of Washington publishes a Vision and Values statement,⁶ which, if taken at face value, seems to align with the goals of this project: where it exhorts the “celebration of place” and lists “respect” among its enshrined values. The study of the ancestral language of the people who have lived for many generations on the land where the University campus now stands falls in line both with paying respect to those cultures and the people who embody them as well as the celebration of the place where they live. Similarly, both the mission statement of the University of Washington⁷ and the Vision and Values statement invoke the embracing of diversity and multiculturalism. As stated by Hale (1992) and placed into focus by Hugo (2010), the vitality of indigenous languages and cultures are connected to attitudes about their values. The *Łəbšucid* project aligns with these goals by shifting the efforts of study and analysis which are too often reserved for the languages of the economically powerful onto the local indigenous language of the place where the University now stands.

In sum, by bringing computational resources to bear on linguistic questions about endangered languages, this project provides an opportunity to improve linguistic science, to participate in the building of technological aides for the documentation and conservation of endangered languages, and to take a small step away from the traditional ethnocentric marginalizing of the indigenous languages and cultures of the Pacific Northwest. Having outlined the motivations, in the next section I provide a more technical presentation of the project’s main research question and how that question is addressed within this work.

1.2 About the author

To introduce oneself in Lushootseed requires one to state not only their name, but where they come from — often who their parents are and possibly where they came from. Since my project seeks to honor Lushootseed language and culture, I believe it’s important for me to introduce

⁶<https://www.washington.edu/about/visionvalues/>

⁷<http://www.washington.edu/admin/rules/policies/BRG/RP5.html>

myself carefully before embarking on further presentation.

I was born in the city of Virginia Beach, approximately two miles inland from the Atlantic Ocean. My mother and her family are from central North Carolina. My father's family comes from Wytheville, in Virginia mountains, but he grew up in Suffolk, near Virginia Beach. I grew up there, at the mouth of the Chesapeake Bay, until I was a young adult. Then I spent several years teaching English lessons to adults in cities outside of the United States. In learning to teach English, I was introduced obliquely to linguistics and became enamored with the International Phonetic Alphabet. One year after moving to Seattle in 2006, I became a non-matriculated student in the Department of Linguistics at the University of Washington; this is how I began to acquire a formal education in linguistics.

From the time I was 17 years old to the time I was 26 years old, I had lived in 7 cities in 3 countries. Some of this movement was due to unhappiness with the political and cultural conditions of the American southeast. Some of this was due to the restlessness of youth and the fortunate circumstances which having an English language teaching certification and concomitant work experience afforded me for travel. Nevertheless, very shortly after arriving in Seattle, I felt that my movements had ended, that I would stop moving and grow roots.

For me, growing roots meant learning all I could about local history and culture and ecology. When I began to read materials I found in bookshops and libraries, I was disappointed to find that stories almost always began with and were centered on European explorers or American settlers. Those stories are interesting and valuable to learn from, but the cheerful ignorance they often displayed towards the existing peoples, their place names, their stories struck me as glib at best, and racist at worst.

From 2012 to 2016 I lived in tultx^w (Carnation, Washington). Shortly after moving there, I was admitted to the PhD program at the UW and I knew by then that I wanted to study Lushootseed. I visited the Canoe Family language and culture classes on the invitation of Angela *yu?yu?bæč stadəy?* Wymer. Angela welcomed me to study Lushootseed with her and introduced me into a series of online classes with Zalmi *?əswəli* Zahir. My heart overflows with thankfulness towards Angela, because she opened a door for me to a world of insight and expe-

rience — the deep history I was looking for. As a linguist, then, the best I can hope for is that I can honor those who have shared their language and culture with me by endeavoring to do something good, something helpful.

Vi *taq^wšəblu* Hilbert used to tell a story from Klickitat wherein a girl who is strange or inadequate (*?əsqcab*) is looked upon with compassion by a wise cedar tree. The cedar tree instructs the girl to take his roots and prepare them and then to twist them into a basket. She is told to place a design on the basket. Upon completing the task, she is instructed to take the basket to the river and dip water and bring it back. When she returns, there is very little water remaining in the basket. She is told that her work wasn't tight enough, to do it again. Four times the girl makes a basket with a design. On the fourth time, the basket is full of water upon returning to the tree: the girl has made something beautiful and useful. She is instructed to take it to the oldest person in her village and give it away.

In my mind I relate to the strange girl in a certain way: I have begun something which I hope can be useful and even perhaps beautiful, but more directly, right now I know that what I'm offering here doesn't yet hold water. This dissertation document describes a first basket, one from which lessons can be learned.

This is my introduction, then: I have found a home on Puget Sound. I am a student of Lushootseed. I want to be helpful.

1.3 Introduction to the Dissertation

Having presented myself, and the motivation for my endeavor, I now return to the task of introducing it. As I mentioned above, the form of the main contribution here is a series of case-studies in Lushootseed Grammar Engineering. These studies show how Lushootseed exhibits an import set of data which raises certain challenges for some standardly assumed models of syntax and semantics. These challenges are sometimes technical ones, which is especially true in the case study regarding valence-increasing affixes (Chapter 7), where the facts of Lushootseed lead me to implement a unique hybrid joining Parson's "subatomic" semantics (Parsons, 1990) with the more classically-Davidsonian traditional approaches of Minimal Recursion Semantics (Copes-

take et al., 2005). I argue for a principled and limited violation of DELPH-IN's⁸ Characteristic Variable Principle in order to maintain a more constrained semantic space. In other cases, the challenges are more theoretical, as seen in the case study which explores noun phrases and non-verbal predicates (Chapter 8). In that study, I try to put a finer point on the classic question of whether there is a noun-verb distinction in Salish by implementing two competing analyses and comparing them in the light of a testsuite of Lushootseed sentences. These two systems are then further compared in an evaluation exercise which uses them to analyze the four traditional stories published at the end of Lushootseed Reader volume I (Hess, 1995). In addition to the comparison of the two approaches to non-verbal predicates, the evaluation also captures the state of the work to-date and the error analysis therein provides a roadmap of what's next for grammar development.

The structure of this dissertation is as follows. Chapter 2 provides an introduction to the Lushootseed language: where it is spoken, how its distribution was affected by colonization, how it relates to other languages of the region and some typological features of note. This is followed by an introduction to the linguistic theory and framework which I assume: Chapter 3 introduces Finite State Grammar (Johnson, 1972; Kaplan and Kay, 1994; Beesley and Karttunen, 2003), which undergirds my approach to morphophonology; Head-Driven Phrase Structure Grammar (Pollard and Sag, 1994), which provides the overall theory of syntax; DELPH-IN's description language for implemented grammars (Copestake, 2002), in which the actual linguistic rules and constraints are written; and Minimal Recursion Semantics (Copestake et al., 2005), which is used to describe semantics within this framework. Those concepts and frameworks connect my work to the broader linguistic community but the day-to-day tools and development environment also form an important part of the context of this project, so these are introduced in Chapter 4. After this introductory material, I begin to present the project development. Chapter 5 presents my initial implementation of a morphophonological analyzer which can map Lushootseed text into a morpheme-regularized form appropriate for a syntactic grammar. After that, in Chapter

⁸<https://delph-in.net/>. DELPH-IN is introduced in this document in §3.4.

6, I present the two smaller case studies: one treating the general word order of Lushootseed with a focus on the main-clause subject markers which seek sentence-second position, the other describing the integration of Lushootseed object-marking (a morphological phenomenon) with that word-order treatment. These smaller case-studies lead to the longer-form case studies of valence-increasing morphology (Chapter 7) and of non-verbal predicates (Chapter 8). The evaluation and error-analysis are presented in Chapter 9. The last chapter provides a summary and conclusion.

The last remaining matter for this introductory chapter is to list-out a number of idiosyncrasies which pertain to the glossing and referencing of materials in this work.

- Linguistic glosses and translations are retained from the original where they exist, but added when the original is unglossed, or partially glossed. When the glosses are mine, I follow the morphological classification of Beck (n.d.) unless otherwise noted, modulo the following point.
- Beck's glossing of Lushootseed examples provides detailed information about determiner types. However, since the particular subtypes of determiner do not factor into the points I attempt to illustrate in this document, I have usually simplified the determiner glosses to just `DET`.
- The full list of abbreviations used in linguistic glosses is presented in Table 1.1. These are taken from the recommendations in the Leipzig Glossing Rules (Bickel et al., 2008), where provided there. The glosses themselves are almost wholly adopted from those of Beck and Hess (2014, 2015).
- Most of the linguistic examples were published in Beck and Hess 2014 or Beck and Hess 2015. Sometimes these examples appeared first in previously published materials. When the example can be found in Beck and Hess 2014 or Beck and Hess 2015 I cite the example by the English title of the story as it appears in Beck and Hess 2014 or Beck and Hess 2015

followed by the initials of the storyteller in parenthesis. For example, the Lushootseed text⁹ of example (1) was first published in Hess (1995) and appears in Beck and Hess (2014) as the first line of “Mink and Tutyika as told by Edward Sam”. I hope that this style of citation will help readers to connect to the particular context of the example in a way which a standard in-text citation with only author, year and a page number would not. The names and initials of speakers who are referred to in this way are listed here:

- (AJ) *lalacut* Agnes James, Snohomish
- (AW) Alice Williams, Upper Skagit
- (DM) Dewey Mitchell, Upper Skagit
- (DS) Dora Soloman
- (ES) Edward “Hagan” *s?adacut* Sam
- (HM) Harry Moses
- (LA) Louise Anderson
- (ML) Martha *səswixab* Lamont
- (MS) Martin Sampson

- (1) *tuχ^wax^w čəd łuyəhubtubicid ti tusyəhub ?ə tu'di?*
 tuχ^w=ax^w čəd łu=yəhub-tx^w-bucid ti tu=s=yəhub ?ə tudi?
 just=now 1SG.SBJ IRR=recite-ECS-2SG.OBJ DET PST=NMLZ=recite OBL DET
tusluχ^lluχ^l čəł
 tu=s=luχ^l-luχ^l čəł
 PST=NMLZ=DSTR-old 1SG.POSS

‘I am simply going to recite to you a traditional story from our distant Elders.’

Mink and Tutyika (ES)

⁹See §2.4.1 for an introduction to Lushootseed orthography and its correspondence with pronunciation.

Having introduced my project, my motivations and the structural prospectus, the next chapter continues the introduction by turning to the Lushootseed language itself.

1	first person	IMP	imperative
2	second person	INCH	inchoative
3	third person	INTJ	interjection
ACT	causative of activity	INTNS	intensivity
ADD	additive	IRR	irrealis
ADNM	adjunctive nominalizer	LOC	locative
ALTV	allative applicative	MD	middle
ATTN	attenuative	NEG	negative
AUTO	autonomous activity	NMLZ	nominalizer
CAUS	causative	OBJ	object
CNTRFG	centrifugal	OBL	oblique
CNTRPT	contrapuntal	PASS	passive
CONJ	conjunction	PFV	perfective
CONN	connector	PL	plural
CONT	continuous	POSS	possessive
CSMD	causative middle	PR	preposition
CTD	contained	PRIV	privative
DAT	dative applicative	PRLV	prolative
DC	diminished control	PROG	progressive
DEM	demonstrative	PROX	proximal
DET	determiner	PRTV	partitive
DIM.EFF	diminished effectiveness	PST	past
DIST	distal	Q	question particle
DSTR	distributive	QTV	quotative
ECS	external causative	SBJ	subject
F	feminine	SBJV	subjunctive
FOC	focus	SCONJ	sentential conjunction
HAB	habitual	SG	singular
HMN	human	STAT	stative
ICS	internal causative		

Table 1.1: Abbreviations for linguistic glosses used in this document

Chapter 2

LUSHOOTSEED

This chapter provides a brief introduction to the Lushootseed language: its endonymns and areal distribution, its history, its philogenetics and some notable typological features.

2.1 Language Name and Areal Distribution

2.1.1 Lushootseed Names and Dialects

Lushootseed is the English name for the indigenous language of Puget Sound. Historically, in English, Lushootseed has been referred to as Puget Salish, or, just as often, writers named the language in local terms: referring to Nisqually, Snoqualmie or Snohomish. In fact these local languages are all part of a network of intercommunicative dialects which span the Eastern Puget Sound from Bellingham Bay to the Nisqually River delta — including the islands of the Sound and some areas of the Kitsap Peninsula (see §2.1.2).

The language’s endonym(s) (*dx^wləšucid/tx^wəlšucid/x^wəlšucid*) vary along with dialectical and idiomatic usage across the region (Zahir, 2016). The root morpheme of the word Lushootseed *ləš* refers directly to the Puget Sound region (Bates et al., 1994, vii). The word contains a prefix *dx^w-/tx^w-/x^w-* (“contained”) and a suffix *-ucid* (“mouth”), which, when used in concert, indicate “language” so the analytically the word Lushootseed may be translated as “the language of the Puget Sound region” (Bates et al., 1994, vii).¹

Linguists subdivide the Lushootseed dialect continuum into two main groups, Northern and Southern, along a border which runs East to West roughly from the location where the

¹The idea that the “language” reading of *-ucid* requires the co-occurrence of *dx^w-* “contained”, is disputed by Zahir (pc), who points out the form *pastəducid* (“English”), from *pastəd* (“Boston”) + *-ucid*, which does not bear the contained prefix. Zahir posits that the occurrence of the contained prefix on *dx^wləšucid* indicates the inherence of the language to the region indicated by *ləš* (pc).

Skykomish River meets the Snoqualmie River (near Monroe, Washington) to the shore of the Puget Sound just North of Edmonds, Washington. Snohomish, Skagit and other distinctions north of this border are grouped as Northern Lushootseed. Duwamish, Snoqualmie, Puyallup and others south are grouped as Southern Lushootseed. The Northern and Southern Lushootseed dialect groups cluster together according to morphological, lexical, and phonological features (Hess, 1977).

One of those phonological features is reflected in the prefix which forms the first part of endonym from which the English word “Lushootseed” is sourced. So, within Lushootseed-speaking communities, the indigenous word for Lushootseed is attested in three forms: *dx^wləʃu-cid*, *tx^wəlʃucid*, *x^wəlʃuʔcid*. These names are in rough correspondence with the dialect divisions drawn by linguists: the first of these refers (roughly) to Northern Lushootseed, the second to Southern Lushootseed, the third to a Lushootseed spoken farther south around Nisqually.

Historically, Lushootseed speaking was centered around the winter village sites located along transit corridors formed by the rivers (Suttles and Lane, 1990). Each of these villages would have presented dialect variation with respect to other villages, had idiosyncratic words, etc. Nowadays, after the imposition of the reservations and the dispossession of land (Marino, 1990), the migration to reservations has destroyed smaller differences which would have been found from village to village and river to river (Hess, 1977).

2.1.2 *Lushootseed Speaking Area and Puget Sound Place Names*

Lushootseed speaking territory is surrounded on three sides by related Salish languages. It is bordered on the North by *Lhéchalosem* (“Nooksack” [nok]), spoken around the body of water now known as Bellingham Bay. To the South, Lushootseed is bordered by speakers of *Q^wayáyilq* (“Upper Chehalis” [cjh]), spoken South and West of Nisqually. Across the Sound to the West, Lushootseed is bordered by *tuwaduqutSid* (“Twana” or “Skokomish”, [twa]) of the Eastern Olympic Peninsula. Lushootseed is bordered on the East by the crests of the Cascade mountains (Hess, 1977). See the map in Figure 2.1 for approximate language boundaries.

Lushootseed has provided the source for many English place names around Puget Sound —

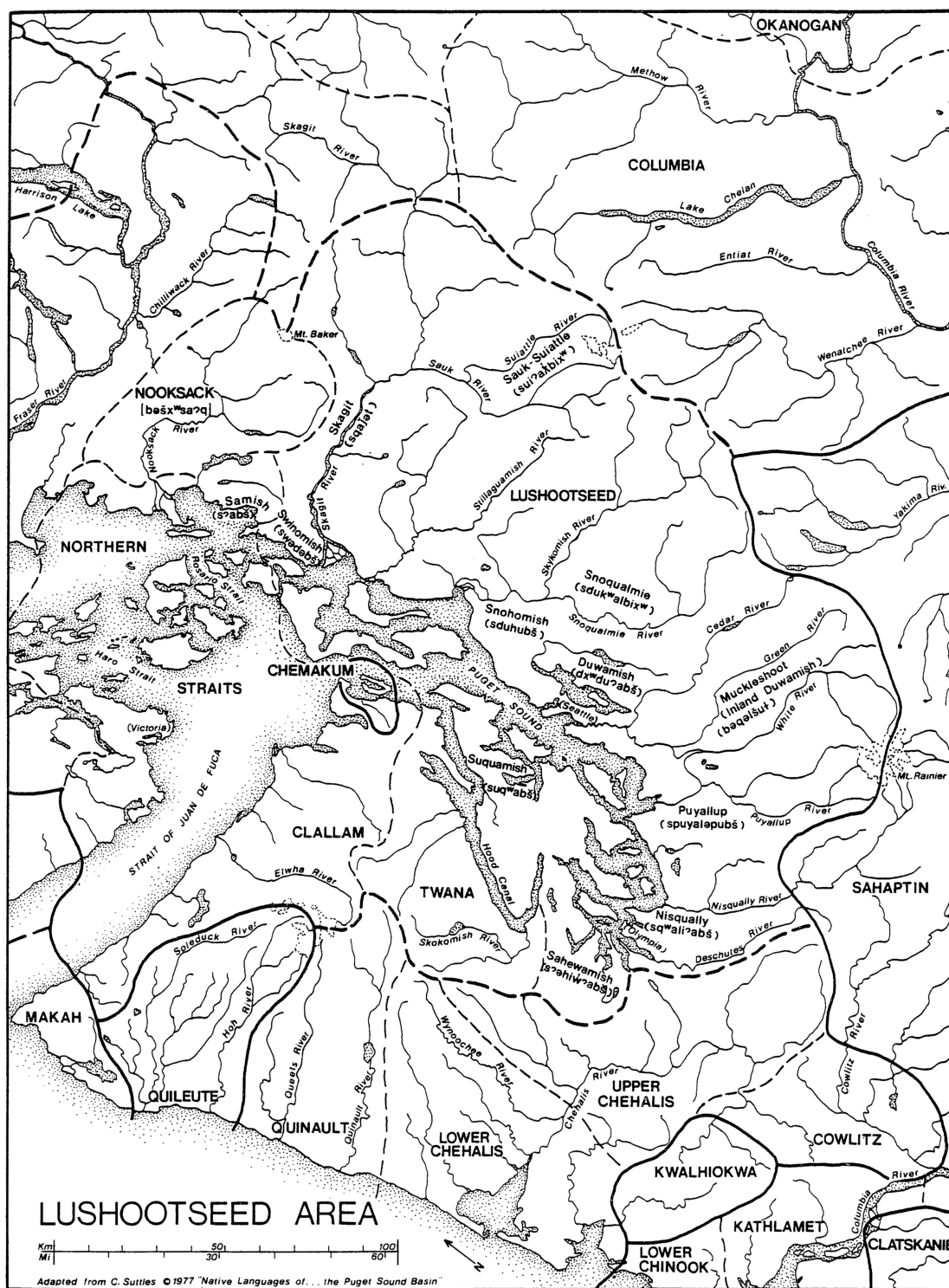


Figure 2.1: Lushootseed speaking area in geographic and philogenetic context. The solid lines represent divisions between language families. The darker dashed lines represent divisions of subgroupings within Salish. The lighter dashed lines represent language differences within a subgroup.

(Modern) Lushootseed	English
liqtəd	Licton
šilšul	Shilshole
sbaq ^w abqs	Mee-kwa-mooks
sduhubš	Snohomish
sduk ^w albix ^w	Snoqualmie
swəpəbš	Swinomish
bəqəsuł	Muckleshoot
puyalup	Puyallup
tultx ^w	Tolt
dx ^w sq ^w ali	Nisqually

Table 2.1: Selected Lushootseed place names and their English corollaries. Mostly from (Hilbert et al., 2001)

from and parks and neighborhoods such as Seattle’s Licton Springs (From Lushootseed *liqtəd* “red”, referring to the ferrous rocks around the spring which were used for dyes) and West Seattle’s Mee-kwa-mooks park²), to cities, rivers, and towns all over the region. Table 2.1 provides a small selection of English place names and their modern Lushootseed sources. In fact, one plot twist on these pairings is that because of a phonological shift which took place in Lushootseed since the place names were first borrowed, the English place names of many locations record a facet of an older version of Lushootseed: the existence of nasal stops (see §2.4.1).

The historical record of Lushootseed place names is especially indebted to T. T. Waterman, who undertook a survey of Seattle area place names from local informants during his time at the University of Washington in the early years of the 20th century (Waterman, 1922), and to

²I hypothesize that the name “Mee-kwa-mooks” is from *sbaq^wabqs*, the Lushootseed place name for Alki point (Hilbert et al., 2001, pgs. 59, 61, 66). Despite the fact that the city of Seattle, at the time of this writing, has published literature and websites claiming (without reference or citation) that the name of Mee-kwa-mooks park, near Alki point, comes from a Nisqually word meaning “shaped like a bear’s head”(City of Seattle, 2017). No Lushootseed speakers or scholars that I consulted with have been able to connect this claim to any relevant Lushootseed vocabulary. Because there is no source cited on the City’s literature, I am unable to ascertain the origin of this unlikely hypothesis. On the other hand, the Lushootseed word for Alki Point *sbaq^wabqs* is very similarly shaped — especially under consideration that Lushootseed nasals became voiced stops in the mid to late 1800s.

Vi *taq^wšəblu* Hilbert, Jay Miller and Zalmi *ʔəswəli* Zahir, who annotated and expanded the Waterman manuscript (Hilbert et al., 2001), providing maps, photographs of (many of) Waterman’s informants, translations of the place names into English (where possible), etymological notes as well as a transliteration into modern Lushootseed orthography. This revision of Waterman’s work forms the basis of much of the Lushootseed which appears in the Burke Museum’s *Waterlines Project* maps (Sheikh and Collins, 2009).

2.2 *History*

2.2.1 *Early history*

Speculative prehistory (Thompson and Kinkade, 1990, 45–46) based on the distribution of related languages and other evidence suggests that Proto-Salishan speaking peoples slowly spread out from an original settlement on the Fraser River delta, eventually covering the entire territory now known as the Salish Sea³ and inland to the Rocky Mountains. Using linguistic evidence in concert with ecological facts suggests that the Salish homeland may have extended as far South as the Skagit river (Kinkade, 1990), into what became Lushootseed speaking territory. While Kuipers (2002) demurs about time-depths of linguistic branchings and sound changes,⁴ archaeological researchers have suggested that the expansion outward from this Coastal *Urheimat* may have begun approximately 2000 years ago (Carlson, 1990, pg. 107). This means that Lushootseed, or an early form of it, has been spoken in the Puget Sound region for thousands of years.

Early accounts of Lushootseed within the written record begin with Gibbs (Hess, 1996). Gibbs was a lawyer with political connections to the Whig party who had served as the Collector of the port of Astoria during the Fillmore administration and had subsequently travelled north to Puget Sound, settling for some years near Fort Steilacoom (Stevens, 1873, 13). During his time in what was then Washington Territory, he collected vocabularies, place names, and

³The Salish Sea refers to the protected salt waters of Puget Sound, the Straights of Juan de Fuca and Georgia and the smaller and bays and canals which connect them (U.S. Board on Geographic Names Domestic Names Committee, 2009).

⁴“Of prehistoric migrations and fusions of Salish (and possibly non-Salish) groups we know nothing beyond what the languages themselves can tell us.” (Kuipers, 2002, pg. 2)

compiled bilingual dictionaries of several local languages: Lushootseed (Gibbs, 1877), Clallam and Lummi (Gibbs, 1863a), and Chinook Jargon (Gibbs, 1863b). Gibbs' facility with indigenous language and culture, as well as his education as a lawyer led to his inclusion in the treaty negotiations carried out in 1855–1856 between Washington territorial governor Isaac Stevens and the indigenous peoples. Gibbs drafted the language of the treaties. Gibbs' legacy in the matter of treaties is interesting in that even though the negotiations were not carried out in good faith by the American side (Reddic and Collins, 2005)⁵ and were directly responsible for sickness, suffering and death,⁶ ultimately, the language which Gibbs included regarding the guarantee of hunting and fishing rights at “usual and accustomed places” had important and specific ramifications in *United States v. Washington* in 1975, the outcome of which is known popularly as the Boldt decision.⁷ Miller (2016) connects the inclusion of that language to Gibbs' sympathetic observation of the reservation conditions in Northern California.

In all, Gibbs lived in the Pacific Northwest for only a decade, but left an outsized legacy with respect to early documentation of Lushootseed: Gibbs' Nisqually dictionary contains approximately 3000 Lushootseed items with English glosses followed by a smaller section of English to Lushootseed. In addition to his dictionary, it is Gibbs' spellings of Lushootseed words as placenames which were reified into English (Miller, 2016).

Another figure who provided early documentation of Lushootseed was Eugene Chirouse. Chirouse was a Catholic missionary, noted for his ability in Lushootseed (Riddle, 2009). Chirouse penned an instructional pamphlet with 14 lessons on Lushootseed structure (Chirouse, 1880). I wasn't able to obtain a copy of this book by the time of this writing, so I'm not able to

⁵The character of the American position in these negotiations is documented, *inter alia*, in Reddic and Collins (2005). Of note in Isaac Stevens' own writings is the message from Isaac Stevens to the Legislative Assembly on 28 Feb 1854 (Gates, 1940, 3) and a letter to the Commissioner of Indian Affairs George Maypenny (11 May 1855 WSIA, R.1) which reveals the intent to unilaterally change the terms of the agreement after the signing:

...it is believed that as soon as the central agency shall be underweigh [sic], all the special reservations can be dispensed with and the Indians consolidated on the general reservation [at Tulalip].

⁶See, *inter alia*, Chapter 2 of Dover (2013).

⁷*United States v. Washington*, 384 F. Supp. 312 (W.D. Wash. 1974), *aff'd*, 520 F.2d 676 (9th Cir. 1975)

compare Chirouse's transcription and analysis to the dictionary of Gibbs. However, as far as I can tell, Chirouse's writings contain the second-oldest Lushootseed in the historical record.

2.2.2 Assimilation

This history of Lushootseed cannot be summarized without mention of the assimilation period. From the early 17th century until the Meriam report of 1928, United States had placed indigenous people and their culture under direct attack (McDonnell, 1991; Williams, 2013). The attack was carried out officially through racist laws and disingenuous treaties (Reddic and Collins, 2005) as well as unofficially, by turning a blind eye to murder and other crimes against indigenous people (Kluger, 2011; Dover, 2013). Even the treatment of disease was employed in such a way as to further the purposes of assimilation (Pearson, 2003). For generations, the Lushootseed people faced an ultimatum from a powerful enemy: they could abandon their language, culture and ancestors and assimilate into Christian capitalist agriculturalism, or they would be destroyed. The position of the United States, as reflected in the policies and laws of the government as well as the writings, attitudes and actions of its citizens, came from a mindset of Social Darwinism and Manifest Destiny in which European culture and civilization was seen as the apex of a hierarchy, with all other cultures seen as inferior and destined to evolve towards the European model or die out. These beliefs provided a rationalization for acts and beliefs which were inhumane and racist (McDonnell, 1991; Marino, 1990).

For the most part, indigenous people of Puget Sound were spared the brunt of these policies until the 19th century simply because there were no large white settlements in the area. Vancouver's sea-probe of the area took place in 1792. The overland arrival to the mouth of the Columbia by the Americans party led by Lewis and Clark would not take place until 1805. The first permanent settlement of white colonists in Puget Sound would not take place until the founding of Nisqually House in 1832. Even this was really just an outpost along the road the "Cowlitz portage", an intermediary point between Fort Vancouver on the Columbia and Fort Langley on the Fraser (Bagley, 1915).

The establishment of Nisqually House via the Hudson Bay Company was the United King-

dom's attempt to create a foothold in the latently-disputed territory which stood between Great Britain's claims to the North and Spain's claim to the south. The land which was then known as Oregon Territory had been established by the Anglo-American Convention of 1818 (Bevans, 1968) as a zone which neither the United States nor Great Britain could claim to own nor setup a government of, but that the citizens of either nation would be allowed to settle and trade and travel there. This agreement was known by both sides to be a temporary one, an affirmation that the facts on the ground (military presence and settlements) would determine the eventual owner (Kluger, 2011, 22–23).

The matter of control of the Pacific Northwest was resolved 1848, 30 years after the original 1818 agreement, when the United States and Great Britain agreed that the border along the 49th parallel would be extended to the Sea and then along the Strait of Juan de Fuca, with a proviso that the Hudson Bay Company could continue to do business as usual and retain its property in the United States until the US Government offered it compensation. In some sense, this is when the assimilation policies of the United States towards indigenous peoples began to be officially applied within Puget Sound.

As recounted in Meeker (1905), the indigenous people of Puget Sound exhibited hospitality towards the newcomers to the region. But as more and more settlers arrived, encroaching on resources and ways of life, tensions developed. When crimes were committed by whites against natives, officials looked the other way (Dover, 2013). When natives were accused of crimes against whites military search parties and capital punishment were effected against the accused (Kluger, 2011). Assimilation policies, some of which are outlined below, worked to destroy native culture and language as white settlement continued unabated. While ultimately, assimilation policy failed in that indigenous culture is alive in Puget Sound today, its impacts on human lives and civilization were substantial.

Treaties

Having gained control of Oregon Territory 1848 (from an international perspective), United States law began to have an official effect on the residents of Puget Sound. In 1850, the United

States Congress passed a law which would exacerbate settlement pressures on traditional lands. The DONATION LAND ACT provided title to 320 acres for single men and 640 acres for married couples who arrived in Oregon Territory before the end of 1850 (and half as much to those who arrived before 1854) as long as they lived on and cultivated the land for four years (United States Government Printing Office, 1850). This act was legally questionable, however, because existing United States law also made it clear that no one was allowed to take indigenous land without the permission of the affected people:

Article III. Religion, morality, and knowledge, being necessary to good government and the happiness of mankind, schools and the means of education shall forever be encouraged. The utmost good faith shall always be observed towards the Indians; their lands and property shall never be taken from them without their consent; and, in their property, rights, and liberty, they shall never be invaded or disturbed, unless in just and lawful wars authorized by Congress; but laws founded in justice and humanity, shall from time to time be made for preventing wrongs being done to them, and for preserving peace and friendship with them. (Northwest Ordinance of 1787)

Other US Laws known as the “Trade and Intercourse Acts” ensured that only the United States federal government was empowered to make offers to purchase indigenous land (Kluger, 2011, 28). This led to an acrimonious situation in which whites had motivation to take land by any means, and the US Government was motivated to enact treaties with the indigenous people quickly, in order to normalize the situation (Wilma, 2003).

In 1853, Washington Territory was split off from Oregon Territory and Isaac Stevens was made its first governor. Stevens also simultaneously became its first superintendent of Indian affairs. Stevens set out to impose treaties which would form a legal basis upon which the United States could rationalize the transfer of title from indigenous people to settlers. Stevens saw the tribes as an impediment to civilization: his job was to extinguish Indian title to land in order to facilitate settler development (Rasmussen, 2014; Marino, 1990). To this end, he carried out

four treaty making sessions with natives West of the Cascades during the three months between December of 1854 and February of 1855. At each of the sessions some gifts were presented, the treaty was read out in Chinook Jargon, some commentary was allowed, then the leaders⁸ were invited to step forward and place their mark on the treaty (Wilma, 2003).

Dover (2013) characterizes the Indian perspective on the treaty of Point Elliot. She tells of a private council between the Indians in which they discussed their options, and recounts the speeches of three individuals:⁹

Why should we sign anything? Don't sign anything. That piece of paper is the white man's piece of paper. Let the white man sign his paper; that is his paper. Let the white man sign it, because we don't know what it says. The white man will tell us what it says, but we don't know what it says. Don't sign anything. Let's kill them all. Why should we let them live? They are going to change everything. We don't have to listen to them.

–Sq^wəšəb¹⁰ (of the Snoqualmie Tribe) (Dover, 2013, 22)

Why not sign the treaty? We are not going to be allowed to go home. You see what we have to go through every day. There are soldiers here, and they are here with their guns. They come here, get off of that ship and then stand around here with their guns. We have to sign that treaty — we should. We better sign in. Then we can all go home. Later on, perhaps, we can talk about this again. Maybe there will be a different white people we can talk to. There is no use trying to protest and create trouble. There is no choice. There is no choice.

⁸The “leaders” were particular individuals who had been selected by Steven’s representatives as the “chiefs” of the Indians whose land he wished to take (Reddic and Collins, 2005).

⁹Typically, I have followed the convention of using standard Lushootseed orthography and *italic* font on Lushootseed words and names which appear in the running text of this document. Capital letters are not used in standard Lushootseed orthography. The names of the individuals below are represented in this text as they appear in Dover 2013, however: with capital letters and without italics.

¹⁰Sq^wəšəb’s name appears on the treaty as “Smoke” — which is a valid translation for this name — but Dover emphasizes that the name would have been more properly translated as “Fog” or “Mist”.

–Tyee William (Dover’s Father’s Uncle), known on the treaty as Stishail (Dover, 2013, 22)

Tell the white man they are welcome to live on this land as long as they want to. We will sign a paper where we will allow them to use our land and then they can just pay us so much for a year. Someday in the future we will talk about it again, but to sign something and make it final — that we will sell the land and we take the reservation — we should not do that.

–Tləbšičtəd, known in English as Club Shelton (Dover’s Father’s Great Uncle) (Dover, 2013, 22)

Dover recounts this meeting and these speeches to emphasize the difficulty of the situation — and the fact that not everyone agreed on the right course of action. Tləbšičtəd, taking two or three others with him, left the negotiations and did not sign the treaty.

The year after the signing of the treaty of Point Elliot, indigenous people of Northern and Central Puget Sound were required to move to Tulalip, where reservation land had been set aside. Dover (2013, 35) recounts the fact that the Lushootseed people showed up at the reservation site the next year, expecting the promised schools, healthcare and lodging, but instead found that nothing which had been promised was in fact available — no one informed the Indians that the treaties they had signed had not been ratified by Congress (and wouldn’t be for another 5 years).

Culture Outlawed

Dover (2013, 35–46) lays out the initial difficulties encountered by the those who were relocated to the Tulalip reservation: principally, these amounted to the lack of fresh water and food. During certain periods, residents of Tulalip were not allowed to leave, even to fish or hunt. Instead of moving freely along fresh water streams and living communally in long houses, people were expected to construct single family houses on lots and port water by hand from nearby sources. The indigenous people were told to go and live like the whites but were not afforded the requisite infrastructure. People got sick and died due to the unsanitary conditions and lack of food.

In addition to the difficulties of living on minimal water and food, Indian culture was under explicit attack. Since 1819's "Civilization Fund Act",¹¹ the United States provided funding for religious organizations who provided schools for Indians. At these schools, children were subjected to a militaresque life where they were marched from room to room, forced to do physical labor every day for most of the day, and Indian language was punished corporally. Dover, who attended the boarding school at Tulalip describes in first person the treatment of children who spoke in their native language:

I was given a whipping for speaking our own language in school when I was nine years old. It sounded worse than if I had killed forty people. Every time I think about it, it makes me mad: to have other people babied around and me beat up. I wasn't very big when I was nine years old. (Dover, 2013, 118)

Dover characterizes the effect of this abuse on the culture and succeeding generations:

Believe me, we never talked "Indian" at the school again. Some of our people, such as those who were my sons Wayne's or William's age and some of the girls, were kind of shocked and disturbed over the fact that they never heard our language. They said their mother never spoke the language or their father or their grandmothers, and so they say they didn't know a single word of their Indian language. I gave a talk somewhere to a group, and I explained the reason why we seldom spoke Indian: it was beaten out of us. We were severely punished, and some of the boys and girls got worse punishment than I did. (Dover, 2013, 119)

Dover also reflects on the psychological effect of having been personally denigrated and denied human legitimacy:

My grandmother only talked the Indian language. In the school they spoke nothing but English, and the teachers were all white people. They told us the order was,

¹¹<http://legisworks.org/congress/15/session-2/chap-85.pdf>

you can't speak the Indian language. I don't know if it is really sadness — when you are trying to put those things together, you feel sadness, since you are trying to believe the Indian is important. It's mine, but I have to take the white man's language. I have to take his religion, and I have to reconcile it with what my grandmother said. I realize there are some white people who are kind, but there are others who want to send us off somewhere. (Dover, 2013, 188)

The physical abuse of children who spoke their native languages in Indian boarding schools was not only the policy at the Tulalip school. The policy was widespread, as evidenced in the literature on boarding schools.¹²

Attendance at schools was not optional. In 1893, the 53rd congress of the United States, in appropriating Federal monies to be used to support the treaties they had entered into with indigenous groups, included a proviso allowing the withholding of rations from families who refused to put their children into the boarding schools:

Title 25 of the United States Code §283¹³ reads in part:

The Secretary of the Interior may in his discretion, establish such regulations as will prevent the issuing of rations or the furnishing of subsistence either in money or in kind to the head of any Indian family for or on account of any Indian child or children between the ages of eight and twenty-one years who shall not have attended school during the preceding year in accordance with such regulations. This provision shall not apply to reservations or part of reservations where sufficient school facilities have not been furnished nor until full notice of such regulations shall have been given to the Indians to be affected thereby.

The amount and value of subsistence so withheld shall be credited to the tribe or tribes from whom the same is withheld, to be issued and paid when in the judgment

¹²For an account of some other schools in Washington, see Chalcraft (2004).

¹³Enacted Mar. 3, 1893, ch. 209, §1, 27 Stat. 628, 635., the text quoted here is current U. S. law, in effect, at the time of this writing. <https://www.govinfo.gov/app/details/USCODE-2017-title25/USCODE-2017-title25-chap7-sec283>

of the Secretary of the Interior they shall have fully complied with such regulations. The Secretary of the Interior may in his discretion withhold rations, clothing and other annuities from Indian parents or guardians who refuse or neglect to send and keep their children of proper school age in some school a reasonable portion of the year.

13 years later, Congress allowed for the removal of Indian children from their families to “reform schools”, specifically stating that the consent of the parents or guardians of the children was not required. The following is a direct quote:¹⁴

§302. Indian Reform School; rules and regulations; consent of parents to placing youth in reform school

The Commissioner of Indian Affairs, under the direction of the Secretary of the Interior, is authorized and directed to select and designate some one of the schools or other institution herein specifically provided for as an “Indian Reform School”, and to make all needful rules and regulations for its conduct, and the placing of Indian youth therein: Provided, That the appropriation for collection and transportation, and so forth, of pupils, and the specific appropriation for such school so selected shall be available for its support and maintenance: Provided further, That the consent of parents, guardians, or next of kin shall not be required to place Indian youth in said school. (25 USC §302)

Indigenous families were not only under attack within schools, the “Code of Indian Offenses”¹⁵ specifically forbids particular traditional practices on reservations. The code makes

¹⁴Enacted June 21, 1906, ch. 3504, 34 Stat. 328., the quoted text is current U. S. law, in effect, at the time of this writing, as can be verified by reading the text as published by the United States government at this URL: <https://www.govinfo.gov/app/details/USCODE-2017-title25/USCODE-2017-title25-chap7-sec302>

¹⁵Robert Clinton was able to locate the original text of the Code of Indian offenses with “difficulty”, and subsequently made a transcription available: <https://rclinton.files.wordpress.com/2007/11/code-of-indian-offenses.pdf>

See Chalcraft (2004) for an account of the usage of this law to suppress religious activity in Puget Sound.

specific references to traditional feasts, dances and medicine and the penalty for participating in a traditional dance is the withholding of food. The Code of Indian Offenses led to the Courts of Indian Offenses and Indian Police and was part of the United States government's systematic attack on indigenous culture — attempting to supplant traditional structures with Western, colonial ones (Clinton, 2007).

It's hard to fix an exact date for the end of the assimilation era, because boarding schools and other assimilationist policies continued late into the twentieth century — and, as mentioned above, in some cases assimilationist policy is still the law of the land. Nevertheless, by the 1920s, certain events mark the beginnings of a shift of policy. The Indian Citizenship Act of 1924 provided that American born Indians were citizens of the United States. The Meriam Report, “The Problem of Indian Administration” was submitted to the secretary of the Interior in 1928 (Meriam et al., 1928). The report strongly criticized the policies of the United States towards the Indians and found that the government was failing at its goals of protection of indigenous people, “their land, and their resources, both personal and cultural”. This report is seen to have influenced the Indian Reorganization Act of 1934 which rolled back the allotment policies which had forced the breaking up of tribal land into individual family plots, allowed the tribes to form governments and create communal land trusts (Prucha, 1984, Ch. 31).

In all, the policies of the assimilation era did huge damage to the health and well-being of Lushootseed people and their culture. However, in the case of the Lushootseed people, the story is one of resilience in the face of this adversity, not one of defeat. Dover (2013, 167–171, 199–229) describes the tribal organizations in Tulalip which were dedicated to improving conditions. Later, in Tulalip, other organizations were created in order to collectively manage resources and culture. She notes the re-inauguration of the return of the first Salmon ceremony in 1982 as a turning point in cultural revitalization. Nowadays, many of the tribal organizations in Puget Sound host language programs dedicated to Lushootseed teaching and propagation.¹⁶

There is a large body of literature on United States assimilation policies and the ramifica-

¹⁶Some discussion of the status of indigenous languages in Washington can be found in Hugo (2019).

tions on Indian lives and culture. This section barely scratches the surface of the topic. I have included it here because in introducing Lushootseed, I wanted to put the decline Lushootseed usage during the assimilation era into some context, and to provide some links to further reading for interested parties. In this exposition, I have drawn heavily on Dover's account of the history of Tulalip, and in that regard, what I have cited here is not necessarily reflective of the experiences of all of the Lushootseed-speaking people who lived through the assimilation era. Lushootseed speaking peoples were gathered into other reservations around the region and the conditions in those places and the experiences of those people are discussed in other works. Then, of course, other indigenous families did not go to the reservations at all. While each person's experience was of course unique, the policies and culture which were prevalent across the region created undeniably adverse conditions for those who wished to preserve and propagate indigenous language. So this era is a part of the history of the Lushootseed language. In the next section, I discuss the history of Lushootseed study in the formal context of written historical record and linguistic analysis.

2.2.3 Twentieth Century Anthropology and Linguistics

Beyond George Gibbs, Eugene Chirouse and T. T. Waterman, I am unaware of further reports on Lushootseed language or culture from the 19th century. By the beginning of the 20th century, the local history of Washington was beginning to be seriously documented by Edmond S. Meany at the University of Washington, in whose materials some occasional written Lushootseed appears.¹⁷ Similarly, Lushootseed words make appearances in Anthropological works such as Haeberlin and Gunther (1930) and Gunther (1945). Full focus on Lushootseed language structures would not appear until the second half of the twentieth century, but layman scholars such as Arthur Ballard recognized the value of the content of Lushootseed stories much earlier. The foundational materials upon which much of modern scholarly study of Lushootseed rests were recorded by Leon Metcalf, like Ballard, a layman who took it upon himself to make recordings

¹⁷See, for example, the appendix of Hunt and Kaylor (1917), where some Lushootseed placenames appear with translations.

of elder speakers. In this subsection, I provide brief profiles of these contributors.

Arthur Ballard

Arthur C. Ballard was born to a settler family in Slaughter (later Auburn) in 1876. Ballard was interested in languages from a young age: at 15 he had begun collecting a Yakima word list. He went on to study Latin, Greek, German, Spanish, and Esperanto (Wilma, 2000). Ballard accompanied and aided T. T. Waterman in his fieldwork on Puget Sound Indians, his facility with the local culture and languages earning him high praise from T. T. Waterman (Watson, 1999; Payton, 2000).

In contrast to other researchers of his day, Ballard was known to pay his informants for their time and knowledge (Watson, 1999). Several of Ballard's collections were published during his lifetime by the University of Washington. While Ballard's published articles present not the Lushootseed itself, but English translations of Lushootseed legends (with the exception of certain names and songs, which he carefully rendered in the American Phonetic Alphabet), he made wax cylinder recordings of the informants speaking in their native tongue (Ballard, 1929, 33). The Ballard recordings are the earliest audio recordings of Lushootseed.

Because Ballard was careful about recording the names and lineages of his informants, he was able to give back to the community in 1957, when his testimony to the Indian Claims commission provided verification of lineages, locations of villages and resource sites using information he had collected from individuals who were no longer alive (Ballard, 1929; Watson, 1999).

Leon Metcalf

Another repository of audio recordings of Lushootseed are due to the work of Leon Metcalf, a high school music teacher with an interest in languages who noted the disappearance of Lushootseed and foresaw the value of recording elders. In 1950, Metcalf took a language transcription class from Melville Jacobs. Metcalf ignored Jacob's advice that any recording must be transcribed

immediately in order to be useful, and that any transcriber must spend at least three years transcribing with a fountain pen into a notebook before recordings could be made. He bought a tape recorder for \$250 of his own money and began recording elders at Tulalip. He later made recordings in the South at Tacoma and Muckleshoot. In some of his work, he recorded messages between homebound elders who hadn't seen each other in years (siastənu, 2005, 63–65).

Much of Metcalf's recordings were transcribed, translated and annotated by Vi Hilbert in the late 20th century (Hilbert, 1974). Because of Metcalf's foresight, anyone can now hear the masterful storytelling of Martha Lamont, and the history of Puget Sound Indians and whites directly from siastənu (Ruth Sehome Shelton, Harriet Shelton Dover's grandmother).

2.2.4 *Linguistic Study Era*

In the twentieth century, Lushootseed itself became the focus of some scholarly literature. In the 1940s, Melville Jacobs, who had been a student of Franz Boas, was in the Anthropology Department at the University of Washington and began to point students towards the study of the local Salish language. The first few papers to be published which were treating the language's structures directly were short sketches (Ransom, 1945; Tweddell, 1950). Eventually, Thom Hess took up the mantle, and his work on Lushootseed laid a foundation upon which much subsequent scholarship on the language was built. This is especially true not only because Hess made recordings of elder native speakers and transcribed their words into a standardized spelling system which he helped to create, but also because he is credited with the inspiring Vi *taq^wšəblu* Hilbert to begin her work on Lushootseed. Hilbert was a native speaker of Northern Lushootseed who would spend the next 50 years working on the conservation, documentation and propagation of Lushootseed. Hilbert, as both a scholar and a speaker, was widely recognized as extraordinary: she was named a living treasure of the State of Washington (Yoder, 2004).

The first Lushootseed publication to emerge from Jacob's tutelage was Ransom (1945), which lays out a grammatical sketch of Duwamish: treating the phonetics and phonology, some morphophonological rules, basic syntax patterns, the numeral distinction between animate and inanimate and an initial exploration of the directional system. The next was Tweddell (1950), a

Master's Thesis completed a few years after Ransom's sketch. Tweddell, working with informants based in Seattle and Muckleshoot, continued Ransom's start on a descriptive grammar of Lushootseed. Tweddell (1950) presents an analysis of 9 classes of prefixes and 13 classes of suffixes in his description of morphology. Tweddell also provides an initial description of the tense and aspect system including the interaction of multiple morphemes within it. Tweddell addresses nominalization, writing out a full verbal paradigm with all of the persons (under his analysis of tense and aspect). His grammar touches on lexical suffixes (see below, §2.4.2). Tweddell's grammar also addresses aspects of the language which interface with other fields of study such as the etymologies of certain loan words and the Lushootseed color system. Tweddell's work with Lushootseed informants was brief, but his work, under the direction of Melville Jacobs, marks an important milestone in the history of Lushootseed description.

Warren Snyder began working with Snoqualmie, Muckleshoot and Suquamish informants around the same time as Tweddell, and made further progress on grammatical description (Snyder, 1957). Snyder's dissertation research was later revised into two books and published as Snyder (1968a,b). The second of these volumes provides a corpus of texts in Snyder's transcription: traditional Lushootseed stories as told by Jerry Kanim (nephew of Patkanim, one of the signers of the treaty of Point Elliott), an autobiographical text by Amelia Sneatlum of Suquamish, and some snippets of conversation between the Sneatlum family. There is also a map with place names from the Suquamish territory. Snyder's work, especially when held in comparison to the preceding from Ransom and Tweddell, pushed forward greatly the amount of documented materials on Lushootseed. Nevertheless, as an anthropologist, his linguistic analysis was faulty in certain ways (Kinkade, 1970), and because he only worked on the language a short time, he never returned to provide any syntactic description of his materials. This would have to wait for Thom Hess.

Hess

Thom Hess began field research on Northern Lushootseed in the fall of 1961 (Hess, 1992). He spent the summers of 1963–1966 working with Harriet Shelton Dover, Elizabeth Krise, Martha

and Levi Lamont, Edward Sam and Louise George. This work led to the presentation of his thesis in 1967, *Snohomish Grammatical Structure*, which provided the first detailed grammar of Northern Lushootseed. Hess's PhD was granted by the newly formed Department of Linguistics at the University of Washington. His work surpassed earlier writings on Lushootseed in terms of analysis, methodology and supporting data. Hess left a lasting contribution on Lushootseed studies: he is credited (by Vi Hilbert) with coining the English name for the language (Zahir, 2016), for standardizing the orthography (Hess, 1984), the principal dictionary (Hess, 1976; Bates et al., 1994), a two part series of pedagogical works (Hess and Hilbert, 1995a,b), the three-part Lushootseed Reader series (increasingly complex texts and audio recordings with accompanying grammatical lessons: Hess (1995, 1998)), and a long bibliography of academic papers.

Throughout his academic career, Hess focused on a few of the languages of the Salish Sea, producing materials which not only advanced professional scholarship but which also facilitated language education in the communities he worked with. In writing about Hess, van Eijk (2009) says that “he generally shied away from publications that would only benefit the academic world, but [...] he concentrated on works that would be useful to the Native language communities where he obtained his data.” Despite all this, it's arguable that Hess's most lasting contribution to Lushootseed came indirectly through his inspiration of Vi *taq^wšəblu* Hilbert.

Hilbert

Vi *taq^wšəblu* Hilbert was born in 1918 on the Upper Skagit river. Her first language was Lushootseed. She would go on to provide much of the basis upon which the Lushootseed language began revitalization in the second half of the twentieth century. Hess introduced Hilbert to linguistic studies and transcription in 1967 when he was working with *čisx^wisał* (Louise George) to analyze a tape-recorded story told by Louise Anderson (Hilbert's mother) (Hess, 1992, 25).

Working with Hess over the next two decades, Hilbert transcribed and published a large body of Lushootseed literature alongside a dictionary, and several pedagogical works. Hilbert's native ear, along with the training in transcription, allowed her to undertake the work of transcribing

the Metcalf tapes (Hess, 1992; Hilbert, 1974, 1976), as well as the Snyder tapes (Hilbert, 1986) and the Ballard collection (Yoder, 2004).

Hilbert not only worked as a collector and transcriber, she was also a teacher. She began teaching Lushootseed classes at the University of Washington in the 1970s. In 1983, Hilbert founded Lushootseed Research¹⁸ a non-profit organization which is dedicated to sustaining research and publications on Lushootseed. This organization has functioned as a publishing house for many of the materials which Hilbert produced: both the translated and transcribed conversations and traditional stories from the Metcalf tapes and also the pedagogical lessons which Hilbert developed for use in her classes. It also continues to operate as an intertribal nexus for students and researchers of the language. Since 2009, Lushootseed Research has hosted an annual conference with presentations and status reports from language revitalization programs around the regions, demonstrations and announcements of how and where to access archival materials, presentations from teachers on tools and techniques.

In addition to her teaching of the language directly, Hilbert instructed people by telling the traditional stories of her people. Hilbert's work was also recognized outside of the Lushootseed community. She was named a "Living Treasure" by the state of Washington in 1989 and was awarded a National Endowment for the Arts Fellowship in 1994 which was presented to her by President Bill Clinton (Yoder, 2004). Within the Lushootseed community, her efforts provided much of the documentation and materials which are available today. Many of her students are now teachers of the language. Her transcriptions and translations form a bridge which connects contemporary speakers and scholars with the recorded words of elders who were alive in the 19th century.

2.2.5 *ICSNL and General Linguistic Study*

Contemporary with the work of Hilbert and Hess, interest in Lushootseed continually broadened in part because a linguistic community grew up in the region which was dedicated to the study

¹⁸<http://lushootseedresearch.org/>

of Salish and Neighboring languages.¹⁹ Since Hess, other scholars have continued the work on Lushootseed grammatical structure and analysis. The archives of the International Conference on Salish and Neighboring Languages²⁰ (ICSNL) provide a lot of materials on Lushootseed grammar by scholars such as Dawn Bates and David Beck (among many others, including Hilbert and Hess). Crisca Bierwort and Toby Langen have done considerable analysis on the narrative style of the traditional Lushootseed stories. Beyond the published papers of professional scholars, many of the tribal organizations which represent the first peoples of Puget Sound also continue to do research on their own materials — to document and analyze the speech of their ancestors as recorded in their own archives.

This account brings the history of Lushootseed description up to the moment, at least, I hope, in terms of milestones. A fuller description of Lushootseed materials in the written records has been compiled in Miller (n.d.). In the next section I describe the relationship of Lushootseed to other languages in the region.

2.3 Philogenetics

Lushootseed is a member of the Salish (or Salishan) language family. Salish is comprised of 22 languages. This section introduces the grouping of Salish languages in order to contextualize Lushootseed within its family. The study of the subgrouping of Salishan languages begins with Boas and Haeberlin (1927), who provide a synopsis of the major sound changes represented in their comparative data collected in expeditions across the region in the late 1800s. Swadesh (1950) applies the method which now bears his name to Boas and colleagues' data and derives the first genetic grouping of Salish (Table 2.2).

Later scholars who have revised this grouping have not made drastic changes. Thompson (1979) moves Bella Coola (Nuxalk, [ISO-639-3: blc]) into its own branch at the top level, grouping the remaining three to level groups from Swadesh into a single “main branch” (Table 2.3). Swadesh' Olympic branch is renamed to Tsamosan and promoted to sister with Coast and In-

¹⁹<https://lingpapers.sites.olt.ubc.ca/introduction-to-the-kinkade-collection-the-on-line-archive-of-ics>

²⁰<https://lingpapers.sites.olt.ubc.ca/icsnl-volumes/>

- I. Bella Coola
- II. Coast Division
 - A. North Georgia Branch
 - 1. Comox
 - 2. Seshalt
 - 3. Pentlatch
 - B. South Georgia Branch
 - 1. Squamish
 - 2. Nanaimo Group
 - i. Fraser
 - ii. Nanaimo
 - 3. Lkungen Group
 - i. Lummi
 - ii. Lkungen
 - iii. Clallam
 - 4. Nootsak
 - C. Puget Sound Branch
 - 1. Skagit-Snohomish
 - 2. Nisqualli
 - D. Hood Canal Branch
 - 1. Twana
- E. Olympic Branch
 - 1. Satsop Group
 - i. Cowlitz
 - ii. Chehalis-Satsop
 - 2. Lower Chehalis
 - 3. Quinault
- III. Oregon Division
 - A. Tillamook
 - B. Siletz
- IV. Interior Division
 - A. Lillooet
 - B. Thompson Group
 - 1. Thompson
 - 2. Shuswap
 - C. Okanagon Group
 - 1. Okanagon-Colville-Sanpoil-Lake
 - 2. Spokane-Kalispel-Pend d'Oreille
 - D. Columbia
 - E. Coeur d'Alène

Table 2.2: Salishan Genealogical Grouping by Swadesh (1950)

- | | |
|--|---|
| <ul style="list-style-type: none"> I. Bella Coola II. Main Body <ul style="list-style-type: none"> A. Coast Division <ul style="list-style-type: none"> 1. Central Salish <ul style="list-style-type: none"> i. Comox ii. Pentlatch iii. Seshelt iv. Squamish v. Nooksack vi. Halkomelem vii. Straights <ul style="list-style-type: none"> a. Clallam b. Northern Straights viii. Lushootseed ix. Twana 2. Tillamook B. Tsamosan Division | <ul style="list-style-type: none"> 1. Inland <ul style="list-style-type: none"> i. Upper Chehalis ii. Cowlitz 2. Maritime <ul style="list-style-type: none"> i. Quinalt ii. Lower Chehalis C. Interior Division <ul style="list-style-type: none"> 1. Northern <ul style="list-style-type: none"> i. Lillooet ii. Thompson iii. Shuswap 2. Southern <ul style="list-style-type: none"> i. Columbian ii. Okanagan iii. Kalispel iv. Coeur d'Alene |
|--|---|

Table 2.3: Salishan Genealogical Grouping by Thompson (1979)

land. Similar changes are exhibited by the grouping of Czaykowska-Higgins and Kinkade (1998) which also flattens out the second-level subgroupings. Kroeber (1999) takes the classification of Czaykowska-Higgins and Kinkade (1998), but adds subgroupings within Central Salish (which he renames “Coast”) for the Straights languages (Clallam and Northern) as well as a subdivision of Tsamosan into Inland and Maritime. Kuipers (2002) adopts the subdivisions provided by Swadesh, but adds a top-level division into Central and Interior Salish, with Tillamook and Bella Coola falling into the Central division. In general, scholars are aware of the subgroupings within Salish but have differed at times as to their exact hierarchy.

One of the major research questions about the subgroups of Salish is whether or not Bella Coola and Tillamook should form isolate subgroups in the far North and far South, or whether they are included in the Central group (Kroeber, 1999). Most scholars (the exception being Kuipers (2002)) segment Bella Coola as a top level isolate. Tillamook, on the other hand, is grouped by Thompson (1979) with Central. Kroeber (1999) follows Czaykowska-Higgins and

Kinkade (1998) in providing a Tillamook group as an isolate, but argues in his text that in certain ways it should probably be grouped with Central.

Lushootseed is situated centrally within the Central (or Coastal) branch of the family. It shares vocabulary and structures with several of its neighboring languages, so research on Lushootseed grammar can inform and provide points of contrast with studies of neighboring languages. Some notable aspects of Lushootseed phenomenology are described in the subsequent section on typology.

2.4 *Typology*

Lushootseed generally presents features which are typical of its family and the Northwest sprachbund, but it also has a handful of notable, unique characteristics for which it holds some fame.

2.4.1 *Phonology*

Tables 2.4 provides the full Lushootseed phonemic inventory.²¹ Most Lushootseed words are structured around monosyllabic roots which are usually CVC or CVCC. However, assuming no underlying schwa (Barthmaier, 1998), there are also CC roots²². Lushootseed reduplication, when targeting such CC roots, yields the potential for many obstruents to occur without intervening vowels. The form, *čʰčʰaʔ* (“stones”), in (2b), illustrates this. Another source of obstruent sequences comes from the fact that a historic final *u* was hardened into *xʷ* in most contexts. The form *dxʷqʷqʷalus* (“white”) (2a) presents the contained prefix *dxʷ*- affixed to a CC root, yielding a sequence of four obstruents. In the broader context of its family, Lushootseed is moderate in its permissiveness of obstruent sequences (Urbanczyk, 1996, Ch. 3).²³

²¹Schwa is included in the chart. In Lushootseed, schwa appears epenthetically and in some reduced contexts. The consensus of the phonologists, however, appears to be that Lushootseed is like other Salishan languages in not having any underlying schwa (Barthmaier, 1998).

²²CC roots are often realized as CɔC but may not be, depending on the sonorance of the C in question and the stress pattern of the surrounding material.

²³Lushootseed’s Northern sister-language, Nuxalk [ISO:639-3:blc] is world-renowned for its obstruent sequences (Urbanczyk, 1996, Ch. 3).

	labial	alveolar	alveo-palatal	lateral	velar	uvular	glottal
plain	p	t c (ts)	č (tʃ)		k k ^w	q q ^w	
ejective	p̣	ṭ c̣ (ts')	č̣ (tʃ')	ɬ' (tʃ)	ḳ ḳ ^w	q̣ q̣ ^w	ʔ
voiced	b	d dz (dz)	j (dʒ)		g g ^w		
fricative		s	ʃ		x x ^w	χ (χ) χ ^w (χ ^w)	h
sonorants			y	l	w		
laryngealized			ỵ	ḷ	ẉ		
			i	u			
			ə				
			a				

Table 2.4: Lushootseed Phonological Inventory given in Lushootseed Orthography. The orthography hews closely to Americanist Phonetic Notation (Bloomfield and Bolling, 1927). In this chart, International Phonetic Alphabet (International Phonetic Association, 1999) equivalents for Lushootseed orthography are given in parentheses where they differ significantly.

- (2) a. *g^wəl dsg^wa?* *ti?iɬ dx^wq^wq̣^walus*
g^wəl d-sg^wa? *ti?iɬ dx^w-q^wq̣^w•alus*
 SCONJ 1SG.PO-ONE'S.OWN DIST CTD-white•colour

‘And the white one would be mine.’ (Beck and Hess, 2015) Star Child (SSP)

- b. *k^wədad čəɬ čɬa ʃədəd dx^w?al ti?ə? sɬ^wuyilcəb*
k^wəda-t čəɬ čɬa ʃəd-t dx^w-?al ti?ə? s=ɬ^wuy•ilc-b
 taken-ICS 1PL.SBJ 1PL.CONJ pushed-ICS CNTRPT-LOC DET NMLZsteam.cook•round-MD
čɬ'čɬ'a?
čɬ'čɬ'a?
 DSTR-stone

‘We have grabbed her and pushed her onto the cooking rocks.’

(Beck and Hess, 2014) Basket Ogress (Agnes James)

Lushootseed is among a group of five Coastal Pacific Northwest languages which are without nasals consonants. The others are tuwaduqutSid [twa] (the only other Salishan language in the

group), Makah [myh] (Wakashan), Ditidah [dtd] (Wakashan) and Quileute [qui] (Chemakuan). All of these languages were recorded as having at least some nasal consonants in the period of early contact with whites, so there is a bit of a mystery in how languages from three families across a non-contiguous area would undergo the same rare sound change at the same time. The hypothesis put forth in Kinkade (1985), drawing on evidence from early descriptions, is that the sounds recorded as nasal consonants during the early contact period were made with a partially open velum, and that these partial-nasals were constant across the entire region. Then, the sound change which resulted in a loss of nasals in five of the region's languages was actually a settling out of this interstitial articulation towards voiced stops and other languages of the regions settled back on nasal consonants. In fact, Lushootseed retains nasals in certain contexts, such as endearing language used towards pets and children, and the speech of Raven (Hess, 1982).

This dissertation contains a few more notes on Lushootseed morphophonology. These are presented in Chapter 5, which discusses Lushootseed phonological structures within the theory of Finite State Grammar (Beesley and Karttunen, 2003). That chapter also describes the methodology and initial results used for constructing an initial morphophonological analyzer.

Connected to Lushootseed phonology is its spelling system. The first spellings of Lushootseed were provided by George Gibbs (see §2.2.1). However, Gibbs' system was idiosyncratic and missed certain contrasts (Zahir, 2009). For linguists and anthropologists, either the International Phonetic Alphabet or the Americanist Phonetic Notation (Bloomfield and Bolling, 1927) provided the necessary tools for documentation, but these alphabets were not taught to Lushootseed speakers. Hess (1984) describes the cultural and linguistic context under which the standard Lushootseed alphabet was created. While mainly following the phonemic principal, Lushootseed spelling uses parenthesis to mark certain phonemes which are dropped or which undergo hiatus in context (allowing readers to easily see paradigms). For the same motivation, certain assimilations are not indicated at all.

The Lushootseed alphabet is important, both practically for learners (as a way to naturally indicate the sounds of Lushootseed without resorting to digraphs or context-dependent rules) as well as culturally. Vi Hilbert described the effect of using the English alphabet to represent

Lushootseed (Hilbert, 2007):

Lushootseed (also called Whulshootseed to indicate the Southern dialect spoken in and around Seattle) is so rich that when it was first written down, the 26 letters of the English alphabet could not catch its 46 sounds. That alphabet is a loose woven basket that won't hold water. Easy ignorance has continued its use in displays at the new MOHAI and for Waterlines at the Burke Museum.²⁴

2.4.2 *Morphology*

Morphologically, Lushootseed has been described as “mildly-polysynthetic” (Beck, 1999, 24) exhibiting multiple forms of reduplication (which can co-occur), valence changing affixes, the lexical suffixes which are a hallmark of the Pacific Northwest sprachbund and a rich array of other affixes which grammaticalize similarity, privation, among others semantic features. Chapter 7 investigates the valence-increasing morphology of Lushootseed, presenting an implemented grammar subsystem which models a subset of Lushootseed combinatoric potential.

Reduplication

Lushootseed reduplication is described in Chapter 5 of Beck (n.d.). Beck's system classifies Lushootseed reduplication into three types based on their forms. Lushootseed reduplication forms can apply generally to the contentful words — whether they be nouns, verbs, or adjectives, with concomitant differences in interpretation.

Most Lushootseed radicals are based around a CVC template.²⁵ Beck's Type I reduplication copies the entire syllable, prefixally. The general meaning of Type I reduplication is associated with distributivity — which may be interpreted as plurality in nouns. In verbs, this marking may indicate that an action or event is distributed across space, or it may indicate iteration. Type II

²⁴The Burke Museum Waterlines Project has amended this error since Hilbert's letter was published. Their current materials use standard Lushootseed orthography.

²⁵As noted above, there are CC, CVVC, CVCC radicals as well. There are particular phonological rules associated with some of these radical types. See Chapter 5 for further detail.

ɬix ^w (three)	buus (four)	
ɬix ^w alʔtx ^w	buusalʔtx ^w	houses
ɬix ^w alg ^w iɬ	buusalg ^w iɬ	four canoes
ɬix ^w alps	buusalps	four domestic animals
ɬix ^w alič	buusalič	four bundles
ɬix ^w əlus	buusəlus	four squares (net), stitches
ɬix ^w ilc	buusilc	four dollars
ɬix ^w qs	buusqs	four points
ɬix ^w ulč	buusulč	four baskets, containers

Table 2.5: Lexical Suffixes used as numeral classifiers (Hess, 1995, 20)

reduplication is associated with attenuativity. It is like Type I in that it is prefixal, but it only copies the syllable onset and nucleus (CV-). The third type is suffixal, -VC, copying the syllable nucleus and coda. This type is associated with several meaning types, the most common of which is usually glossed by Beck as diminished effectiveness (see Beck (n.d., 328) for details). Lushootseed reduplications may co-occur (see Chapter 5). Lushootseed reduplication has also been described within phonological theory by Urbanczyk (1996), which treats Lushootseed reduplication within the phonological framework of Optimality Theory.

Lexical Suffixes

Lushootseed has a set of affixes known as “lexical suffixes”. Lexical suffixes are a hallmark of the Pacific Northwest sprachbund, possibly having spread from Proto-Salishan to Proto-Wakashan via areal contact at a very early period (Fortescue, 2009). These elements are formally bound morphemes with meanings which are interpreted variously as body parts, features of land or buildings, conveyances, etc., depending on context. The usage of lexical suffixes presents a rich system of semantic and syntactic rules. Lexical suffixes are used as numerical classifiers (Table 2.5) according to rules where some lexical suffixes denote shape or functional classes, but other pick out specific types of things (Beck, n.d., 315).

Mithun (1997) argues that lexical affixes cannot be completely analyzed in terms of nominal incorporation or noun-noun compounding.²⁶ Correspondingly, a given lexical suffix can map to a wide variety of English translations, depending on context (3).

(3) a. *ʔas-λi-λp̄•abacəb*

ʔas-λi-λp̄•abac-b

STAT-ATTN-grip.tightly•body-MD

‘Hold blanket around self with hand.’

b. *ʔix^wəɫdat* *ti?iɫ sɔdæg^wabacilsəx^w* *alg^wəʔ ʔə tiəɫ*
ʔix^w•əɫ•dat *ti?iɫ s=dək^wa•abac-il-s=ax^w* *alg^wəʔ ʔə ti?iɫ*

three•time.period•day DET NMLZ=inside•body-INCH-3.POSS=NOW 3.PL OBL DET

čx^wəlu?

čx^wəlu?

whale

‘For three days they were in the body of whale.’

c. *tuləqabac* *ʔə ti?iɫ sčəyus*

tu=ləq•abac *ʔə ti?iɫ sčəyus*

PST=behind•body PR DET head

‘[She] had been shot in the back of the head.’²⁷

d. *put ʔaspil šqabac ti?əʔ hik^w čła?*

put ʔas-pil šq•abac ti?əʔ hik^w čła?

really STAT-flat high•body DET large rock

‘This big boulder is very flat on top.’

²⁶Gerdts (1999, 2003), arguing from Halkomelem, disputes this claim. However, to my reading, the two arguments are not entirely incompatible and may simply be addressing different aspects of lexical suffixes. Gerdts seems to be focussed on the specification of formal, structural analysis while Mithun considers functional criteria and pragmatic criteria as well.

- e. *diʔabac* ʔə tiʔiʔ sčətqs
 diʔ•**abac** ʔə tiʔiʔ sčətqs
 opposite-side•**body** PR DET point

‘On the other side of the point of land.’

- f. *gʷəd•abac* ʔal ti tibu tiʔiʔ spčʉ?
 gʷəd•**abac** ʔal ti tibu tiʔiʔ spčʉ?
 inside•**body** LOC DET table DET basket

‘The basket is under the table.’

- g. *ʔudəgʷabacil* tə dstakəd ʔal tə piit
 ʔu-dəgʷ•**abac**-il tə d=stakəd ʔal tə piit
 PFV-inside•**body**-INCH DET 1SG.POSS=sock LOC DET bed

‘My sock got under the bed.’

- h. *kʷaʔabacəxʷ*
 kʷaʔ•**abac**=axʷ
 released•**body**=now

‘[the bark] let go [of the tree trunk]’

- i. *čitabac*, ʔəʔqʷ**abac**
 čit•**abac**, ʔəʔqʷ•**abac**
 near•**body**, after•**body**

‘Saturday, Monday’

(Hess, 1998, 21-22)

²⁷The Lushootseed orthography and free translation of this item are presented here as they appear in Hess 1998. Of note, however, is that the shooting action referred to in the translation is does not appear in the linguistic forms presented. It seems to me most likely that this example comes from a recorded session with a narrative from which some previous context is being applied. Perhaps the closing bracket in the free translation would be

Referring to water			
as a gap to be crossed		as a means of transport	
diʔucid	across the river	ʔacg^wiɬ	middle of the river
sčəg^wucid	island	sʔilg^wiɬ	shoreline
Referring to a mouth			
as a gap in the face		as a part of the body	
lək^wucid	kiss someone	ʔəsɣɬaʔɬdət	mouth is sore
huyucid	finish eating	təsəbaʔɬdət	lips are chapped
Referring to a door			
as a gap to be filled (a doorway)		as an item which fills a gap (a doorway)	
tqucid	tə šəg ^w ɬ close the door	dx^wʔuq^wyaɣad	open the door

Table 2.6: Referents of Lushootseed lexical suffixes: the same referent may require different suffixes in different contexts (Hess, 1998, 22–23).

In complement to the ability of a single lexical suffix to refer to a multiplicity of objects, a single real-world entity is referred to with varying lexical suffixes depending on the point of view of the speaker. Hess’s examples of this are reproduced in Table 2.6. In the first set of examples *•ucid* is used in reference to a body of water which is a gap to be crossed, but *•g^wiɬ* casts water as a means of transport. In the second set, *•ucid* refers to a mouth as a void or gap in the face (*lək^w*: eat, *huy* make, do, finish), while *•aʔɬdət* denotes a mouth as a body part with subparts (*ɣɬ*: sick, *təsəb*: winter). Finally, a door is a gap or hole in a wall *•ucid*, but it is also an item which fills that gap (*•yaɣad*: arm, wing) (Hess, 1998, 22–23).

Lexical suffixes are an especially interesting phenomenon withing Lushootseed grammar (and, indeed the Pacific Northwest sprachbund). However, for the most part, they fall out of scope of this the aspects of Lushootseed grammar which are under investigation in this thesis. A very rudimentary treatment which I use for certain lexical suffixes is presented in Chapter 9.

better placed after ‘shot’. A free translation of just the sentential material is probably more along the lines of ‘It was on the back of the head.’

ʔalʔal	•alʔtx ^w	house
qədx ^w	•ucid	mouth
čaləs	•ači	hand(s)
bəqsəd	•qs	nose

Table 2.7: Lushootseed Nouns with corresponding lexical suffixes

2.4.3 *Syntax*

Syntactically, Lushootseed is predicate initial, canonically VSO (but VOS is also attested). Because Lushootseed syntax allows nouns, pronouns, adjectives and even prepositions to act as the main predicate of a matrix clause, the term “predicate initial” is more directly descriptive of Lushootseed sentences than VSO, since the main predicate is not always a verb. This particular aspect of Lushootseed syntax is the focus of Chapter 8.

Beyond being predicate initial, Lushootseed exhibits several syntactic elements which target sentence-second (Wackernagel’s) position: the main-clause subject, question-markers and a handful of other grammatical markers. Modeling this word order is the focus of §6.1.

Another notable aspect of Lushootseed syntax is the restriction that no more than a single direct noun phrase may appear as the argument of a predicate. Lushootseed does provide multi-argument constructions, but these constructions will never encode more than a single argument as a direct NP, using instead morphologically marked pronominal reference or oblique prepositional phrases. This system plays a large role in the building of the lexical rules which do object marking, which is discussed in §6.2.1.

2.5 *Conclusion*

This chapter has provided a high-level introduction to the Lushootseed language. I have attempted to touch on location, history, philogenetics and typology. The next chapter turns away from Lushootseed in order to provide an adequate background on the linguistic theory and framework which this dissertation assumes.

Chapter 3

LINGUISTIC THEORY AND FRAMEWORK

In this chapter, I detail the context of this research in terms of the linguistic theory and framework. I start with a very high-level view of the linguistic theory and delve successively downward towards descriptions of the actual representations used in the project.

3.1 Bird's Eye View

In general, this thesis attempts to provide a formalized¹ view of linguistic structures. Wherever possible, I model grammatical phenomena in a mathematically rigorous way. Using computer software to encode the model provides a forcing function for me to achieve that standard. The foundations of such an approach, however, lie in formal language theory. Formal language theory begins by defining a language as a possibly non-finite set of strings. The strings of a particular language are drawn from a collection of symbols called the alphabet of that language. For a given alphabet (traditionally written Σ) and language (L), formal language theory categorizes languages based on the patterns they exhibit. Then, upon these classes, particular methods and algorithms have been developed for deciding whether, for a particular string of symbols w , where $w \in L$ (Hopcroft et al., 2006). In the next section, I briefly discuss the arguments in the literature for treating morphophonology within Finite State Grammar (Beesley and Karttunen, 2003), and in the section which follows, I offer what I believe is the currently accepted rationale for providing a computationally more powerful model for morphosyntax.

Beyond an argument from formal complexity, Bender and Good (2005) examine the practical ramifications — from an engineering standpoint — of nesting a grammar's morphological

¹Bender (p.c.) suggests a distinction between “formalist” (which contrasts with “functionalist”) which holds that an explanation is provided by the formalism and an approach which is “formalized” wherein a formalization is useful in building models and making claims precise.

component inside the syntactic one and vice versa. They conclude, and I concur, that a better approach is to treat the morphological patterns of a language separately from those of morphosyntax, with a shared lexical database containing the information which is pertinent to both subsystems. Such an architecture is congruent with the Lexical Integrity Principle of Bresnan and Mchombo (1995).² More details on the architecture of my project are provided in the Chapter 4.

The linguistic theories at play in this thesis, then, are Finite State Grammar (Beesley and Karttunen, 2003) for capturing morphophonology and Head-Driven Phrase Structure Grammar (Pollard and Sag, 1994) for morphosyntax. In the second half of this chapter I turn from theoretical assumptions to description languages. In §3.4 I introduce the machine-readable format in which the HPSG grammar is written: the DELPH-IN joint-reference formalism, which is also known as Type-Description Language (TDL) (Copestake, 2002), due to its intellectual inheritance from the TDL of Krieger and Schaffer (1994). This language is similar, in many ways, to the formatted attribute-value matrices (AVMs) which are common in HPSG papers. It scans like a linearized, ASCII-fication of such. In many cases, however, it will be convenient to present examples using TDL rather than formatted AVMs. Some of the reasons for this are provided in the subsection below. Additionally, this chapter provides information about the semantic representation language used in this project: Minimal Recursion Semantics (Copestake et al., 2005, MRS). MRS is not a theory of semantics, *per se*,³ but is a description language used for writing constraint-based semantics in HPSG grammars. MRS is designed to accommodate both rigorous semantic theory as well as practical engineering designs when building grammar systems. I introduce the basics of MRS in §3.5.

²See Bruening (2018) for recent argumentation *in contra* Lexical Integrity.

³Some of the theoretical heritage which influenced the design of MRS is given in the Case Study on Valence Increasing Affixes of Chapter 7, in the section on Event-based semantics, §7.2.1.

3.2 *Finite State Morphology and Phonology*

In this section, I present an outline of the basics of Finite State Grammar and formal language theory which motivates the approach taken in this research. This discussion is necessarily abridged. See Beesley and Karttunen (2003) for a full introduction to the linguistic applications of finite state networks.

As mentioned above, formal language theory treats languages as sets of strings. Because the non-finite sets (languages) can never be directly enumerated, one important aspect of the theory is an organization of languages into complexity classes based on the types of algorithms, operations and calculations that can be used to describe them. Broadly speaking, these complexity classes are arranged into the familiar Chomsky-Schützenberger hierarchy (Chomsky, 1959), which is presented in Table 3.1, wherein the formal devices for a describing a language of a type n can be used to describe a language of a type $> n$ but not vice-versa.

Type	Languages	Devices
0	Recursively Enumerable	Turing Machine
1	Context-Sensitive	Linear-bounded non-deterministic Turing Machine
2	Context-Free	Non-deterministic pushdown automaton
3	Regular	Finite state automaton

Table 3.1: The Chomsky-Schützenberger hierarchy of Formal Language Complexity (Chomsky, 1959)

The regular languages are those which are least complex on this hierarchy and therefore, most constrained in terms of the formal apparati which define them. The inductive description in (4) provides the formal definition of regular languages with respect to a particular alphabet Σ (Kleene, 1956).

- (4) a. The empty language \emptyset is a regular language.
 b. $\forall a \in \Sigma, \{a\}$ is a regular language.

- c. If A is a regular language, A^* is a regular language.
- d. If A and B are regular languages:
 - $A \cup B$ is a regular language.
 - $A \cdot B$ is a regular language.
- e. Nothing else is a regular language.

Only three operations serve to generate any regular language from an alphabet: union (\cup), concatenation (\cdot) and Kleene closure ($*$). Union is the same operation as is familiar from set theory. Concatenation of languages A and B refers to a language which contains all strings created by appending a string from B onto a string from A . Thus if $A = \{watch, curtain\}$ and $B = \{ed, ing\}$, then the language $A \cdot B = \{watched, watching, curtained, curtaining\}$. Kleene closure, symbolized by $*$, for a language L indicates the set of strings which contain 0 or more occurrences of any of the strings of L , in any order. These three operations, when applied to regular languages, always define a language which is also regular, and so regular languages are *closed* under concatenation, union, and Kleene closure.

Table 3.1 refers to a finite-state automaton (FSA). An FSA is a theoretical device, which can be thought of as a particular type of directed graph, that can be used to represent a language in terms of acceptance and generation. When an automaton has been written down for a particular language, that automaton can be used to implement an algorithm for deciding whether a given string belongs to the language or not. It is also possible to work from the other direction: systematically exploring the paths through an automaton which lead to accepting states is one method to enumerate the strings of the language represented by the FSA.

Mathematically, a finite state automaton is an n -tuple with 5 components: a set of input symbols (the alphabet of the machine) Σ , a collection of states S , a distinguished “start” state $s_0 \in S$, a transition function δ (which maps a state $\in S$ and a symbol $\in \Sigma$ to a state $\in S$), and a set of final, or accepting states F (Hopcroft et al., 2006). Equivalently, an automaton may be thought of as a graph in which the vertices are states and labeled arcs which connect vertices constitute the transition function. One vertex is the designated start state and 0 or more states

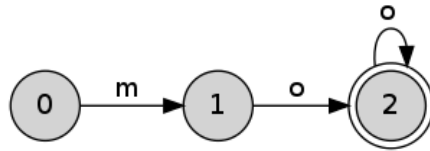


Figure 3.1: Automaton to model the non-finite whimsical language of bovines $/moo^*/$.

will be designated as accepting states. The graph view provides a straightforward way to decide whether an arbitrary string is in the language modeled by the machine. The procedure is to read symbols from the string in question one at a time, moving attention from vertex to vertex on the graph in correspondence with the symbols found on the arc labels.

In more detail, the machine is said to begin in the start state and read the first input symbol. If there is an arc out of the start state which is labeled with that symbol, the machine transitions to the state which the arc points to and then reads the next symbol. If the machine reads a symbol for which there is no arc leaving the current state, it is said that the machine “dies” and rejects the input. No further input needs to be read in order to know that the string of symbols which was input is not a string of the language modeled by the machine. If the machine survives to the end of the input, and if the final state is an accepting state, then the machine accepts the input, i.e. the input is a string of the language modeled by the automaton. To illustrate, consider Figure 3.1, which presents an automaton to model the non-finite language consisting of strings formed by an m followed by 1 or more os .⁴ In this sort of presentation, accepting states are indicated by the double circle around the vertex.

One more preliminary notion, that of a *regular relation*, still needs to be defined in order to apply this background to the domain of interest, the morphophonological properties of Lushootseed. A regular relation is a relation which may hold between two regular languages A and B wherein for each string in A there is corresponding string in B (and vice-versa). Thus a regular relation is a (possibly non-finite) set of ordered pairs (a, b) where $a \in A$ and $b \in B$. The class

⁴The cow language of Figure 3.1 and the Ghost language of Figure 3.2 are really just an extension of the example provided by the Sheep language of Beesley and Karttunen (2003).

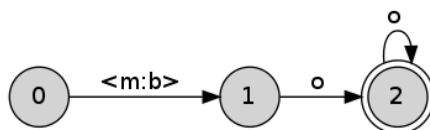


Figure 3.2: Transducer mapping the whimsical language of bovines $/moo^*/$ to a whimsical language of specters $/boo^*/$.

of regular relations are just those which can be modeled by finite state transducers. These are equivalent to the finite state automata discussed above except that (1) the arcs are labeled not with a single symbol, but with symbol pairs and (2) the machine is said to read the first symbol of the pair from the input, and upon taking a transition, to print the second symbol of the pair upon an output tape. In the first usage of this enriched automaton, the machine reads an input from the top side of a tape and prints the output symbol on the lower side in reference to this symbol pairs which label the arcs. Such a machine is said to model the regular relation between two languages, the “upper” and “lower” languages of the machine. Figure 3.2 presents a finite state transducer in which the upper language is the same as the language of Figure 3.1, and the lower is made up of strings constructed by bo followed by any number of os . This regular relation is non-finite: $\{(mo, bo), (moo, boo), (mooo, booo) \dots\}$.

Regular relations or (equivalently) finite state transducers have a closure property which is not relevant for regular languages: that of composition. The composition of two regular relations R and S , written $R \circ S$, is also a regular relation. $R \circ S$ is the relation which maps a string $a \in$ the upper language of R to $b \in$ the lower language of S just in case R maps a to x and S maps x to b . Informally, composition consumes the inner representations. Because regular relations are closed under composition, the composition of two regular relations is also a regular relation, and can also be represented as a finite state transducer mapping an upper language to a lower language.

Finally, the utility of these devices for the construction of a model of morphophonology becomes apparent. I take morphophonology to subsume morphotactics (what morphs can co-

occur and in what order) and phonology (a system mapping canonical forms into surface forms based on their environments). Consider, then, that a finite collection of prefixes is a regular language (call it P) and, similarly, a collection of roots is a regular language (call it R). The definitions and discussion above show that a finite state automaton can be constructed to model these languages. The concatenation of $P \cdot R$ is also a regular language (call it S , a putative language of stems), and can be modeled via a finite state automaton. Crucially, a set of tools to facilitate defining such automata, along with implementations of the operations discussed above, has been released for linguists to take advantage of (Beesley and Karttunen, 2003). Thus, given computer files containing regular languages modelling proclitics, prefixes, roots, suffixes and enclitics we can begin to use these operations to model a system of morphotactics.

Since Johnson (1972), it has been known that SPE-style⁵ phonological descriptions which consist of a series of string-rewriting rules of the form $A \rightarrow B/C_D$, despite their superficial similarity to generalized rewriting systems, are actually employed by phonologists in such a way as to make them equivalent to finite state transducers.⁶ This result is fortunate because the closure properties of finite state transducers mean that, similar to the model given above for morphotactics, well-defined operations can be employed to derive a single regular relation (finite-state transducer) from a cascade of individual transducers. This is especially appealing because the resulting network performs computations which are equivalent to the stepwise rule-by-rule derivations of linear phonology, but without producing intermediate representations.

To provide an illustration of this notion, consider the spelling of the Lushootseed stative prefix cited in (29c). The orthographic form *taslatlil* can be seen as a result of the application of two phonological rules to the canonical form *tu=?as-latlil*. The first deletes the glottal stop of an aspect prefix when it is preceded by a proclitic (5),⁷ feeding the second rule which deletes

⁵*Sound Pattern of English*, Chomsky and Halle (1968).

⁶The crucial point of expressive power lies in a rule's ability to use its own output as a structural description for successive application. Johnson outlines the few phonological rule types which go beyond the complexity of regular relations — they are notably few. See also Kaplan and Kay (1994), who reproduce and augment Johnson's findings using different formal arguments.

⁷All proclitics seem to trigger this, not just ones that end in a vowel.

a *u* preceding a vowel (6).

$$(5) \quad ? \rightarrow \emptyset / _ \left\{ \begin{array}{c} \text{as} \\ \text{u} \end{array} \right\} -$$

$$(6) \quad \text{u} \rightarrow \emptyset / _ (=) \text{V}$$

The rule expressed in rewrite notation in (5) can be modeled as a transducer which captures the relation in which the upper language contains any string and most of those strings are mapped to identical strings in the lower language. However, strings in the upper language which contain a substring which matches the structural description of the rule would be mapped to strings in the lower language in which the structural change has been carried out. This non-finite relation contains pairs as shown in (7). This relation as a finite state transducer is shown in Figure 3.3. The rule in (6), similarly, can be modeled as a regular relation and a transducer can be defined, also shown in Figure 3.3.

$$(7) \quad \{(a, a), (aa, aa), \dots, (ab, ab), \dots, (a = ?as - bb, aasbb), \dots\}$$

Figure 3.4 shows two automata which model regular languages representing the collection of temporal proclitics and two aspect prefixes. In Figure 5, the Kleene closure of the former automaton (representing optional, cyclic application of proclitics), and 0 or 1 occurrences of the latter (representing optional application of an aspect prefix) are concatenated with an automaton modelling the language of the two stems *ʔatlil* and *ʔuχ^w*. This simplistic example with just two stems and two prefixes nevertheless provides a model of a small fragment of Lushootseed morphotactics. Note that any automaton can be also be thought of as a transducer representing the identity relation for that language. Therefore we can use composition to build a new transducer which models the relation of the lexical, or underlying forms, to surface forms with the phonological deletion rules of (5) and (6) applied.

This illustration highlights both the utility of the finite state calculus as a model for implementing and testing hypotheses about the morphophonology of a language as well as the need

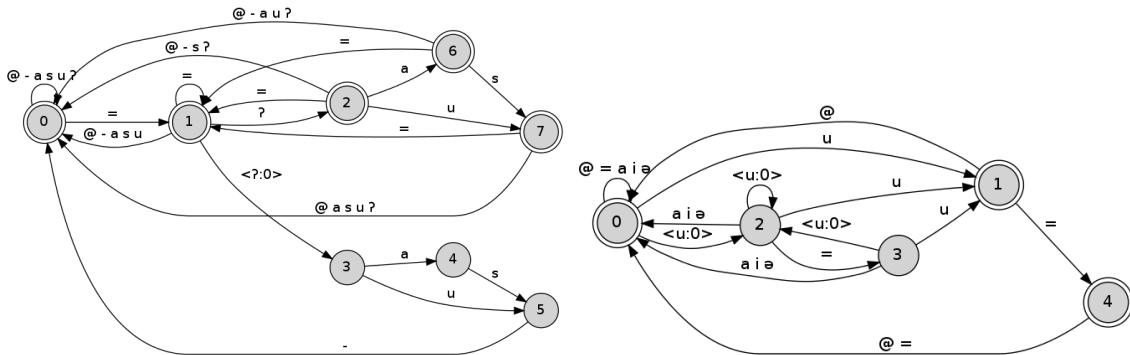


Figure 3.3: Transducers which implement the rules shown in (5) (left) and (6) (right). In these diagrams, the symbol @ is a metacharacter which stands for any symbol (even one not in Σ for a given machine).

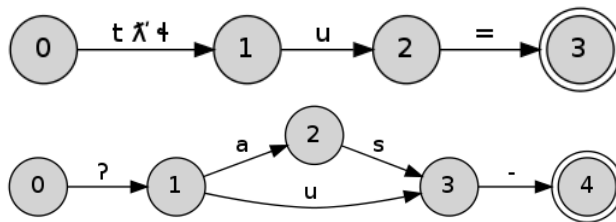


Figure 3.4: Automata to model the language of three temporal proclitics: $\{tu =, \dagger u =, \ddot{x}u =\}$ and the language of two aspect prefixes $\{?as-, ?u-\}$, respectively.

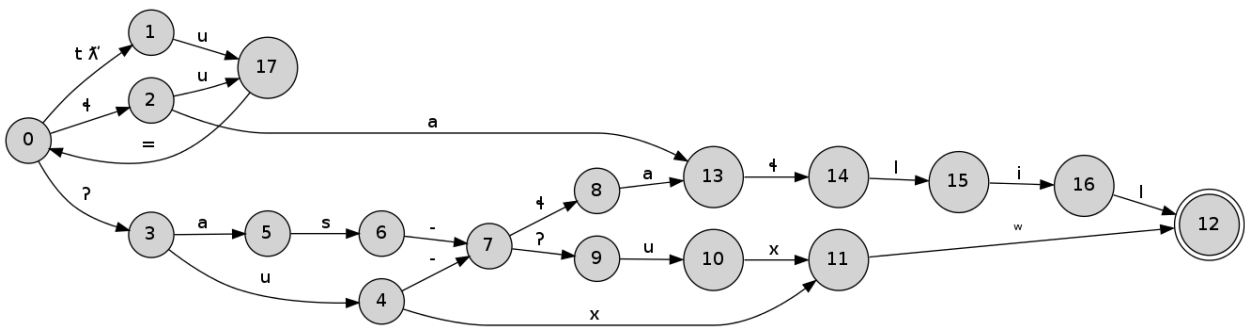


Figure 3.5: Automaton created by concatenating three regular languages: (a) the Kleene closure of the language of Proclitics, (b) the language of 0 or 1 occurrence of the language of Prefixes (both in Figure 3.4) and (c) an automaton which models the language of two stems $\dagger allil$ and $?u\ddot{x}^w$.

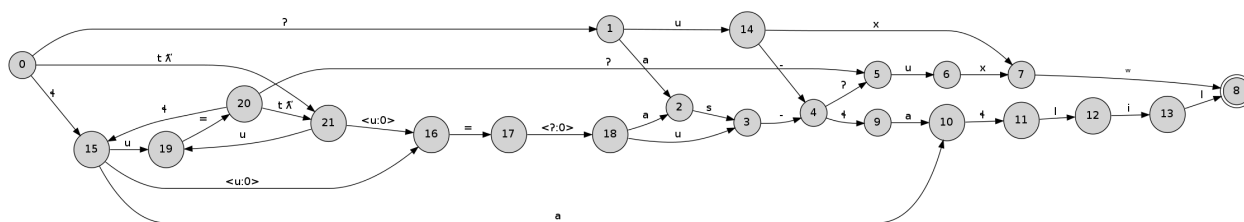


Figure 3.6: Transducer created by composition of the automaton in Figure 3.5 with the transducer of Figure 3.3.

to utilize computational tools to perform the actual operations of composition, and concatenation which are used to build up the network. Figures 3.3–3.6 also highlight the fact that direct manipulation of the networks “by hand” quickly becomes unwieldy. Yet, by defining these networks as a series of concatenated and composed lexicon files and phonological transformations we can take advantage of the perspicuity of the traditional phonological representation without forsaking the utility of compiled, implemented networks which can be used on real data at scale for linguistic hypothesis verification or for the construction of practical language tools.

The preceding discussion illustrates how, within Finite State Grammar, the mathematical definitions of and operations upon finite state transducers are used as a formal basis for implemented models of morphophonology. At a high level, the idea is to first model lexical classes and their combinatory potential as regular languages and operations upon them, yielding a finite state automaton which recognizes strings of the language. After that, one can define phonological rules in a series of finite state transducers, which models the context dependent transformations which map canonical forms to surface variants. Finally, the operation of composition can be used to compile an final, optimized network which models these properties in a single computational object. I implemented this model of morphophonology for Lushootseed in an initial morphophonological analyzer, the construction of which is described in Chapter 5.

3.3 *Head-driven Phrase Structure Grammar for Morphosyntax*

In contrast to the results of Johnson (1972); Kaplan and Kay (1994), the formal complexity of natural language morphosyntax has been shown, despite flawed arguments in the early literature which were pointed out by (Gazdar, 1981, 1983), to be greater than context-free (Shieber, 1985). In fact, researchers in formal language theory, in characterizing the complexity of syntax, have staked out an intermediate ground between context-free and context-sensitive: the class of “mildly context sensitive” languages is used to outline just the sorts of non-context-free constructions and characteristics which occur (Joshi, 1985; Stabler, 2004). In any case, mathematically, Finite State Grammar is inadequate to model morphosyntax, so a more complex model is required.

Prior to the demonstrations by Shieber (1985) that there are natural language constructions which cannot be described by Context Free grammars, Gazdar and colleagues had been working on a framework which sought to remove the transformational component of generative grammar — using meta-rules to capture generalizations across rule classes — allowing the context-free description of particular aspects of language which, in the early transformational literature (Chomsky, 1957, *inter alia*), had been claimed to be beyond context-free. Gazdar’s system was known as Generalized Phrase Structure Grammar (GPSG). However, after Shieber’s proof, the efforts to maintain a purely context-free system had to be relaxed. The intellectual successor to GPSG was known as Head-driven Phrase Structure Grammar (Pollard and Sag, 1987, 1994, HPSG).⁸

This thesis adopts HPSG for at least the following reasons: HPSG is the modern standard-bearer of the original motivations of Generative Grammar (Sag, 2010): it provides a rigorous formalization, implementable in machine-readable form. The descriptive precision of a rigorous formalization is valuable on its face, but in the context of describing natural language, relying on something more precise than natural language itself is all the more important. The

⁸The moniker HPSG has been said to surreptitiously encode its intellectual parentage in the fact that alphabetically, *H* is the successor of *G*. Alternatively, an association has also been made with the fact that the original authors of the HPSG book were contemporarily working at Hewlett-Packard (HP) labs in Palo Alto.

structures, rules and assumptions of HPSG are empirically driven: invisible structures and esoteric movement metaphors are eschewed in favor of information-sharing and a declarative (non-derivational), constraint-based platform. This is not only scientifically parsimonious (in the avoidance of theoretical connundra such as derivational ordering paradoxes and reference to loosely-defined spellout mechanisms), it also provides a foundation which can support the building of practical applications such as parsers and generators in addition to analytical replicability.

HPSG is a formal approach to linguistic theory which can be seen as a continuation of the original principles put forth in formalist writings of the mid-twentieth century carrying forward the emphasis on precise description and mathematical models, but rethinking the apparatus in the light of contemporary methods for managing and manipulating structured information. These techniques are also at the heart of other information engineering endeavors, such as object oriented programming (in software development).

At the core of HPSG is a hierarchy of types which define the space of possible linguistic structures. An HPSG type-hierarchy has a topmost type which is generic, and multiple inheritance allows the factoring of information into relevant subunits, as well as the combination of these subunits into larger ones. For example, Figure 3.7 shows a mini type-hierarchy for modelling some features of animals. This simple example shows how a multiple inheritance type hierarchy factors out relevant groups of information: an animal might be an invertebrate or a vertebrate, but not both (this is captured by the fact that the types *invertebrate* and *vertebrate* have no common descendants); and also allows the combination of these information partitions back into more complex structures: an animal's status as invertebrate or vertebrate is orthogonal to its status as a flyer or swimmer (this captured by the common descendants of each of these pairs).

An HPSG hierarchy not only organizes types into sub/super relationships, any type definition in the hierarchy may also provide feature declarations which define a set of features (and their value-types) which are appropriate for this type and all of its descendants. Feature definitions themselves also make reference to the type hierarchy: the values that a feature can take are constrained according to the types defined within the hierarchy. In this way, HPSG type-

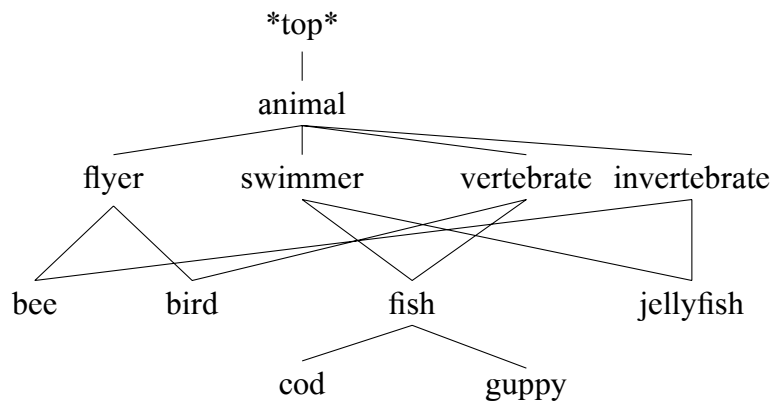


Figure 3.7: A hierarchy of animals. This example hierarchy originally appears in Copestake (2002, 41), and is used here with the author’s permission.

hierarchies are ontologies which provide a complete categorization of everything which exists within the relevant domain. To illustrate how feature definitions work, consider an extension of the animals hierarchy of Figure 3.7 which is presented in Figure 3.8.

In this further articulated example, the feature `WING-COUNT` has been declared on the type *flyer* to take a value of type *number*. The type *number* now appears in the hierarchy as a child of `*top*` with two subtypes: *2* and *4*. In this new hierarchy, all types inheriting from *flyer* will bear the feature `WING-COUNT`. Further, the `WING-COUNT` feature of the type *bee* is constrained to have the value *4*. Note that *bird*’s `WING-COUNT` value is constrained to unify with *number* — just like its parent *flyer*.

HPSG types with associated features like this are referred to as *typed* (alternatively, *sorted*) feature structures and they are often presented as formatted attribute value-matrices (AVM), like the one in (8) — which is the representation of a phrase in Pollard and Sag (1994). Because feature values are types, and types can have features, feature-structures gain complexity through definitional recursion.

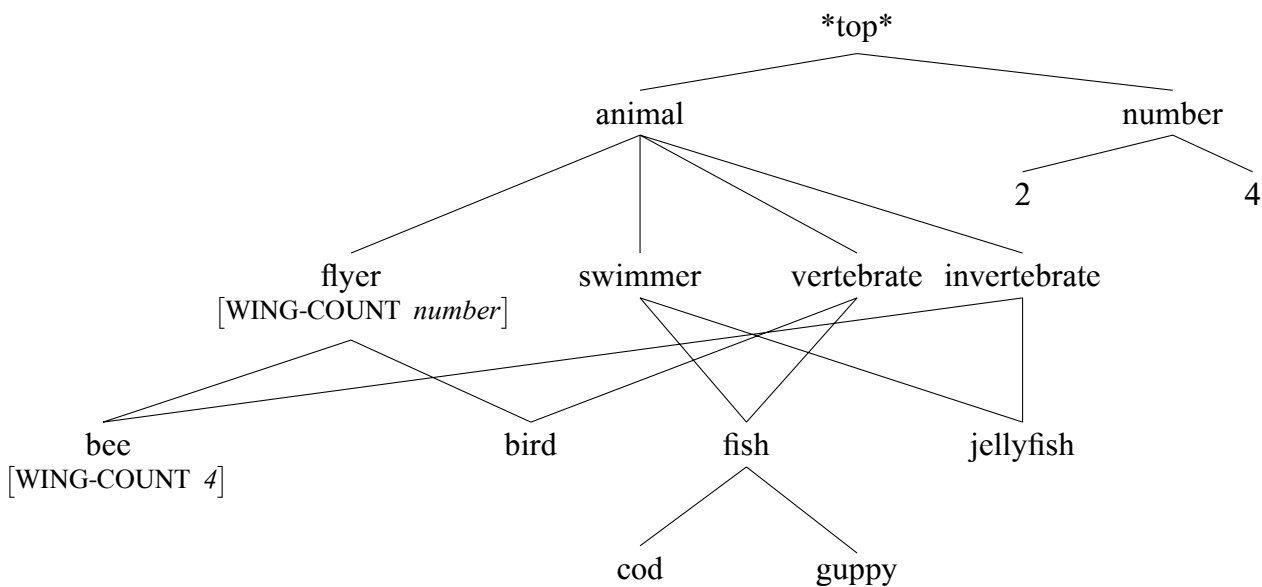
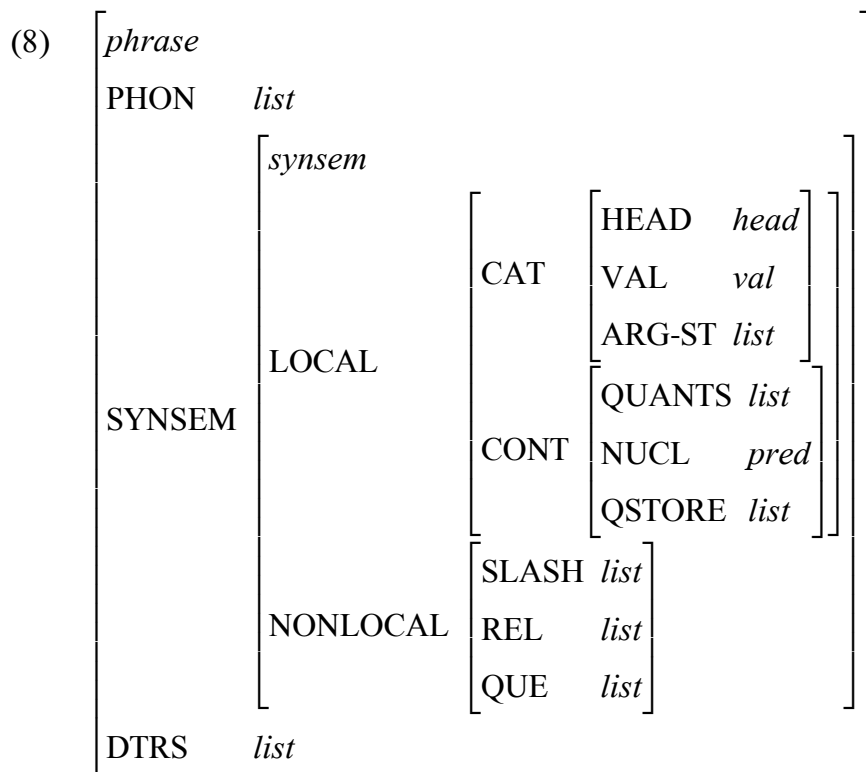


Figure 3.8: An animals hierarchy with feature inheritance and typed feature values.



Another important HPSG data structure — salient in (8) — is lists. Lists and the manipulation of lists are used in the modeling of several grammatical phenomena in a typical HPSG grammar and there are particular aspects of lists which affect how they can be accessed or transformed which merit an introduction. Even in theoretical HPSG, lists are defined as feature structures (Pollard and Sag, 1994, 19) in such a way as to make them directly comparable to the linked-list data structure of computer software.⁹ A non-empty (*ne-list*) list has two features `FIRST` and `REST`. The value of `FIRST` is used to store an item in the list, while the value of `REST` is used to provide a pointer to the next item (9). In order to terminate the list, the value of `REST` is set to the type *e-list*, which unifies with *list* but bears no features.

$$(9) \quad \left[\begin{array}{l} ne\text{-list} \\ \text{FIRST } *top* \\ \text{REST } list \end{array} \right]$$

Linked lists are a simple and elegant way to capture sequential information. However, one ramification of their simplicity is that information on the front of the list is easier to access than information on the end — shortening the list by removing the first element is relatively easy (10),¹⁰ appending to the list by adding an item onto the end is more difficult. If the length of the list is known in advance, one can write a type which will traverse the appropriate feature path (e.g. `FIRST|REST|REST|REST ...`) and construct a new list with the value of the final `REST` having the desired value. However, because HPSG is written declaratively, there are no imperative or control-flow operations to probe a list with something like a `while` loop in order to discover the depth of a list before applying a unification constraint. HPSG types are completely static. One either knows the length of a given list or not.

⁹More acutely, to cons cells (McCarthy, 1960).

¹⁰The type shown here, *list-pop*, is not a usual type in an HPSG grammar, it's a constructed example to show how a list may be shortened. That said, many typical HPSG types do exactly this operation (constructing a shorter list by referencing the `REST` feature of the original). Notable among these are types representing syntax rules which shorten valence lists. Still, in those cases, the feature names are not the ones shown here.

$$(10) \left[\begin{array}{l} \textit{list-pop} \\ \text{SHORTER-LIST } \boxed{1} \\ \text{ORIGINAL-LIST } \left[\text{REST } \boxed{1} \right] \end{array} \right]$$

As it turns out, the majority of linguistic phenomena which are usefully modelled by lists are constrained in such a way that their size is known in advance, and the need for append operations is relatively small. One of the exceptions to that generalization is found in the operations on lists of semantic predicates.¹¹ These lists are built up monotonically via grammatical rule application. Because the semantic composition rules need to be able to append one list onto another, a more articulated data structure is required: the *diff-list*. The `DIFF-LIST` is like a normal list, but it maintains a pointer to the last element. This pointer allows the definition of a generic list append operation which puts two lists together without knowing the length of either one. The types in (11)¹² show the general idea of how a pointer to the last element of a list (`LAST`) can be maintained and then used to join lists of unknown size.

$$(11) \left[\begin{array}{l} \textit{diff-list} \\ \text{LIST list} \\ \text{LAST list} \end{array} \right] \left[\begin{array}{l} \textit{dl-append} \\ \text{APPARG1} \left[\begin{array}{l} \text{LIST } \boxed{1} \\ \text{LAST } \boxed{2} \end{array} \right] \\ \text{APPARG2} \left[\begin{array}{l} \text{LIST } \boxed{2} \\ \text{LAST } \boxed{3} \end{array} \right] \\ \text{RESULT} \left[\begin{array}{l} \text{LIST } \boxed{1} \\ \text{LAST } \boxed{3} \end{array} \right] \end{array} \right]$$

¹¹These comments about lists and their implementation apply specifically to the implemented HPSG framework which I am using in this dissertation. There are versions of “pen and paper” HPSG which take a different approach to list modification and manipulation.

¹²These type definitions are found in the Grammar Matrix core grammar: `matrix.tdl`. The current version under development is available here: `svn://lemur.ling.washington.edu/shared/matrix/trunk/matrix-core/matrix.tdl`. See <http://matrix.ling.washington.edu/> for further details about the source distribution and download instructions.

With that brief overview of HPSG-style type-hierarchies and list manipulation, I can talk about the typical layout of a hierarchy used for grammar engineering. Often, there will be a set of branches at the top which partition the hierarchy into families of information types that interact in a grammatical structure: rules, words, phrases, lexical-entries. Other types will be defined in order to facilitate the definition of basic information bearing structures: very generic types such as booleans or lists are defined in the hierarchy, usually near the top. This is similar to what was illustrated in the animals hierarchy, where the types *swimmer*, *flyer*, *invertebrate*, *vertebrate* were drawn from facts about the phenomena of the animals we wished to model, while the type *number* provided some atomic values useful for feature values. In standard HPSG, the type *sign* is one of the most basic linguistic structure in the hierarchy. *signs* are used to represent words or phrases or even bits of morphology. Signs have features representing their syntactic component, their semantic component and any phonological contribution as well. The type hierarchy also contains types which represent rules for putting together these signs. A rule type will contain a feature structure representing the resultant combination of the daughter elements, as well as feature structures representing the daughters themselves.

The basic operation for combining signs and rules is unification (Kay, 1979; Shieber, 1986; Francez and Wintner, 2012). Unification of two feature structures succeeds when two types are compatible — when one is a more specific subtype of the other (including equivalence) — but fails otherwise. In this way, rule types can constrain their possible daughters to any granularity definable within the type-hierarchy: a rule could be defined to apply to any word, any verb, any verb with a particular feature value, any morpheme, or only a single specific leaf type. An HPSG grammar,¹³ then, is a type hierarchy, along with a specific listing of instantiated types within the hierarchy — the particular rules, phrases, words, lexemes or morphemes in use for a given language. Using such a grammar for parsing or generation also requires the selection of a type, or types from the hierarchy which represent the “complete sentences” or particular

¹³The acronym, HPSG, has come to stand for the grammatical theory, not a particular instance of a grammar built within that theory (although usage varies). So, I will unabashedly write about creating an “HPSG grammar” without fear of the stigmatization associated with expressions like “PIN numbers” or “ATM machines”.

fragment types which one is interested in searching for.

One other important aspect of typed feature structure definitions which bear some introduction is that of “re-entrancy” or co-indexation. This notion refers to the property where the value of a feature at a particular location in a feature structure may be token-identical to the value at any number of other locations: whatever is specified in the first location is equally present in the other. In attribute-value style presentations of typed feature structures, this concept is represented via boxed number identifiers.¹⁴ Like type-hierarchies, typed feature structures can be represented as directed acyclic graphs. The graph view is helpful in understanding the token identity indicated by coindexation in a feature-structure definition: multiple feature paths which lead to the same boxed number indicate multiple ways to walk the graph and end up at the same vertex.¹⁵

With these concepts in mind, I can present some linguistic examples. Just as the example animals hierarchy factored out an animal’s mode of locomotion from the presence (or not) of a spinal cord and then brought these features back together on specific instances, a well-designed HPSG type hierarchy often branches into a number of conceptually independent dimensions representing factors of information about a linguistic entity which are then pulled back together in combination on particular instances. This can be seen in two HPSG types which are used in defining grammatical rules. In (12a), a type which enforces the Head Feature Principal (the principal which says that in a headed rule, the `HEAD` feature of the mother is the same as the `HEAD` feature of the daughter) is presented. Compare this to (12b), a type is presented which copies the `HOOK` information “up” the chain of rule application, from daughter to mother (more on `HOOK` is below).¹⁶ The type in (12c) provides the framework for a generic binary rule which amalgamates information from a distinguished `HEAD-DAUGHTER` and a single dependent. Then,

¹⁴In the memory of a computer, this token-identity is represented via pointers and memory addresses.

¹⁵See Jurafsky and Martin 2009, Ch. 15 for a detailed introduction to features structures and unification algorithms

¹⁶I put scarequotes on *up*, because the mathematical operations which define these structures are order agnostic — so the directional metaphor is potentially misleading or distracting. On the other hand, in a typical scenario, the more specific constraint on the *hook* feature is typically going to be encoded in the grammar on the lexical types, so there is a sense in which this information moves up the tree.

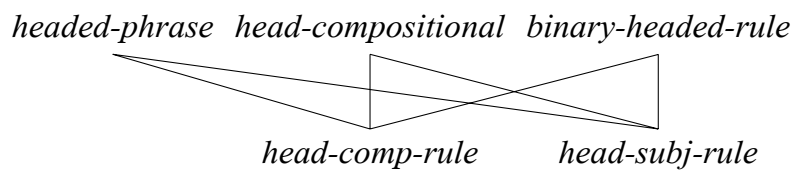
the three types presented in (12a–c) illustrate the breaking down of grammatical phenomena into minimal dimensions which can then be cross-factored in recombinations on particular leaf-type, such as (12d) and (12e) when organized into a type hierarchy such as (13). The independence of a phrase’s headedness from its arity, for example is captured in this sort of model. Then, (14) shows how such *head-subject* and *head-comp* rules might be used within a phrase-structure tree.

- (12) a.
$$\left[\begin{array}{l} \textit{headed-phrase} \\ \text{SYNSEM|LOCAL|CAT|HEAD} \quad \boxed{1} \textit{ head} \\ \text{HEAD-DTR|SYNSEM|LOCAL|CAT|HEAD} \quad \boxed{1} \end{array} \right]$$
- b.
$$\left[\begin{array}{l} \textit{head-compositional} \\ \text{C-CONT|HOOK} \quad \boxed{1} \\ \text{HEAD-DTR|SYNSEM|LOCAL|CONT|HOOK} \quad \boxed{1} \end{array} \right]$$
- c.
$$\left[\begin{array}{l} \textit{binary-headed-rule} \\ \text{SYNSEM} \quad \textit{synsem} \\ \text{HEAD-DTR} \quad \textit{word-or-rule} \\ \text{NON-HEAD-DTR} \quad \textit{word-or-rule} \end{array} \right]$$
- d.
$$\left[\begin{array}{l} \textit{head-subj-rule} \\ \text{SYNSEM|LOCAL|CAT|VAL} \quad \left[\begin{array}{l} \text{SUBJ} \quad \langle \rangle \\ \text{COMPS} \quad \boxed{2} \end{array} \right] \\ \text{HEAD-DTR|SYNSEM|LOCAL|CAT|VAL} \quad \left[\begin{array}{l} \text{SUBJ} \quad \langle \boxed{1} \rangle \\ \text{COMPS} \quad \boxed{2} \end{array} \right] \\ \text{NON-HEAD-DTR|SYNSEM} \quad \boxed{1} \end{array} \right]$$

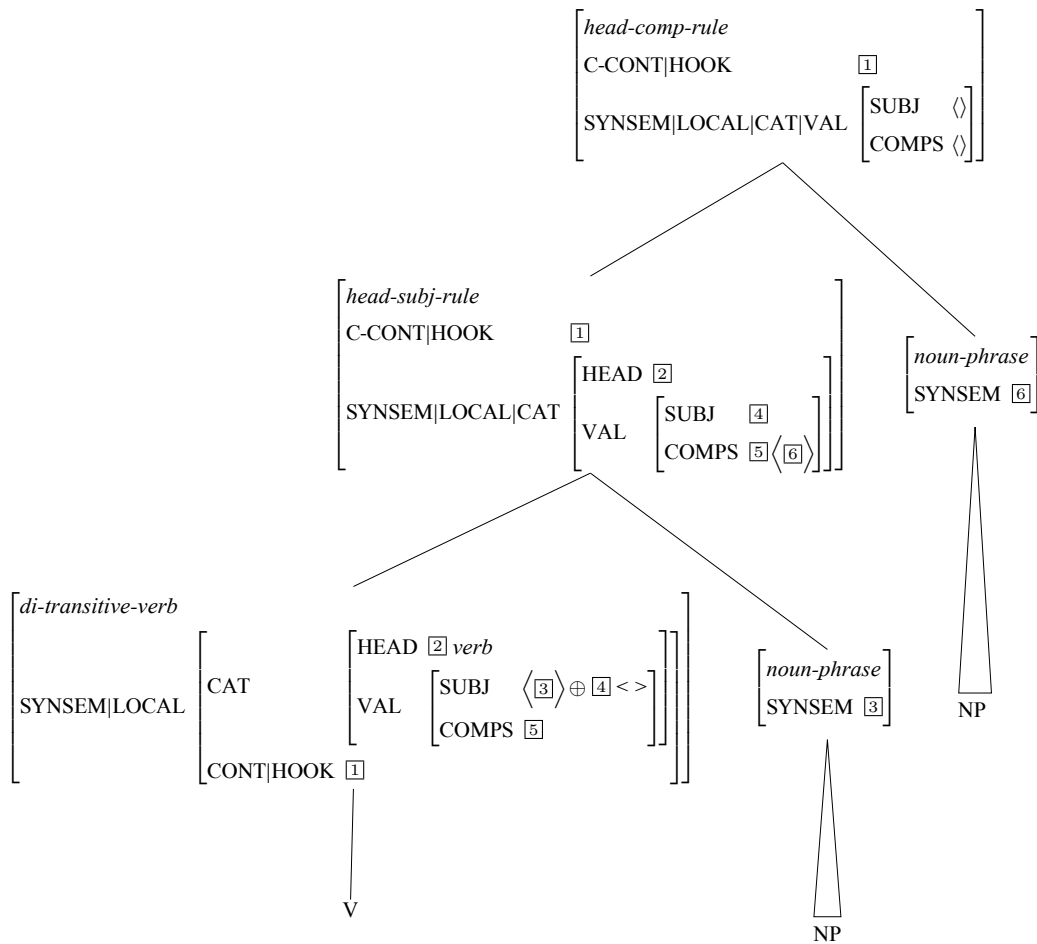
e.

<i>head-comps-rule</i>	
SYNSEM LOCAL CAT VAL	$\left[\begin{array}{l} \text{SUBJ } \boxed{3} \\ \text{COMPS } \boxed{2} \end{array} \right]$
HEAD-DTR SYNSEM LOCAL CAT VAL	$\left[\begin{array}{l} \text{SUBJ } \boxed{3} \\ \text{COMPS } \langle \boxed{1} \rangle \oplus \boxed{2} \textit{ list} \end{array} \right]$
NON-HEAD-DTR SYNSEM	$\boxed{1}$

(13)



(14)



This simple example is meant to illustrate how typed feature structures in a multiple-inheritance context allow a grammar author to take advantage of the factoring out of individual constraints and then their recombination on a particular instance to achieve parsimony of description in the building of linguistic entities. The example in (14) also illustrates another aspect of presenting grammatical types as AVMs: as the information contained within them moves beyond the simplest of toy examples, they quickly become unwieldy to present in their full glory. For this reason, it is traditional when presenting an HPSG analysis to omit information present in an feature structure which is not directly relevant to the matter under discussion. Implemented parsing algorithms will keep track of all the information, however, and will fail to unify structures with conflicting information in parts of the feature structure which is not under discussion. In this way, the use of computer infrastructure to process and define HPSG

structures keeps unintentional errors from creeping into an analysis because the fully specified feature structures are too tedious to manipulate by hand (Bender, 2008).

In this section I have outlined the motivation for using HPSG for morphosyntax in this project. I also introduced some of the basic mechanics involved in the creation of an HPSG grammar. However, one very practical reason for using HPSG is not only its mathematical feasibility for the task at hand, it's the large head-start provided by an existing community of grammarians and engineers who use it. The next section provides further information about my starting point: the tools and frameworks of the DELPH-IN consortium.

3.4 Introduction to DELPH-IN and Type Description Language

In the previous section, I gave a brief introduction to HPSG grammar structures and theoretical apparatus. Because of its precise definitions and mathematical underpinnings, HPSG is amenable to implementation in a computer. In this section I provide an introduction to a particular implementation of HPSG as computer software and the community of researchers and engineers who build and maintain it.¹⁷

Researchers and engineers who use HPSG for modelling linguistic structures in a computational context have recognized the utility of a shared framework and toolkit to advance their work. The DELPH-IN consortium¹⁸ is an affiliation of researchers at sites around the world who have published a large body of tools and implementations all centered on a shared formalism for deep linguistic analysis using HPSG. The tools and implementations of the DELPH-IN consortium are used throughout this project, so there are several elements which must be introduced. Having already provided a bit of introduction to HPSG using the descriptions and formatting typical in linguistic papers, the first element to introduce from DELPH-IN is the machine-readable language used to write out a typed feature-structure hierarchy as grammar code to be consumed

¹⁷The DELPH-IN framework is not the only implemented HPSG framework of note, just the particular one that I have adopted. Among others, the CoreGram project (Müller, 2013, 2015) provides an implemented HPSG framework and has research goals comparable to DELPH-IN's Grammar Matrix (see Chapter 4).

¹⁸<http://www.delph-in.net/>

by a compiler. This language is formally known as the DELPH-IN joint-reference format, but is commonly referred to as TDL (for type-description language), a practice which I follow in this thesis. The grammar of TDL is described fully in Copestake (2002), and in a handful of articles on the DELPH-IN wiki.¹⁹ TDL statements provide a syntax for defining a type using a name, a list of supertypes and any new features (or constraints on the values of inherited features). One of the articles on the DELPH-IN wiki²⁰ shows some basic examples of TDL, a few of which are reproduced in (15).

```
(15) ; Some basic TDL, line-comments are marked with a semi-colon
      type-name := supertype1 & supertype2 &
        [ FEAT1 val1,
          FEAT2 [ FEAT3 val2 ] ].

      another-type := supertype 2 &
        [ FEAT3 #coref & third-type,
          FEAT4 #coref ].

      fourth-type := some-type &
        [ FEAT5 spqr & [ WOMBAT "str" ] ].
```

TDL statements are used to define types (and type-instances) and to place them within a hierarchy. The := operator accomplishes this. To the left is a type name, and to the right a conjunction of supertypes. The last conjunct of this list can optionally be an AVM of features and value constraints. Paths through feature structures can be indicated with the . character (analogous to the usage of the | character in the formatted AVMs found in linguistic papers). Comments are delimited by a leading ;, and continue to the end of a line.²¹ In TDL, co-indexation, or reentrancy, is indicated via identifiers marked with the # symbol.

In many of the examples within this document I will prefer to present HPSG types within this dissertation as TDL statements rather than as the traditionally formatted AVMs. Presenting

¹⁹See, for example, <http://moin.delph-in.net/TdLRfc>.

²⁰<http://moin.delph-in.net/GeFaqTdlSyntax>

²¹TDL authors often follow the LISP convention of indicating the scope of the comment by the number of semi-colons which introduce the remark.

HPSG types using TDL is convenient in this dissertation for three reasons: (1) the actual work of grammar engineering which is presented here uses TDL, so translating to formatted AVMs for the paper version creates a layer of abstraction upon the original, (2) the explicit list of supertypes in a TDL definition statement has no direct corollary in the formatted AVM format (instead, a type hierarchy diagram — like the one in (3.7) — has to be provided) and (3) the mnemonic variable names in TDL coreference notation are easier to digest than the boxed numbers of traditional AVMs.

TDL provides a common description language which DELPH-IN tools can consume in providing functions for parsing and generation. A second, core aspect of the DELPH-IN shared formalism is the usage of Minimal Recursion Semantics (MRS), which is introduced in the next section.

3.5 *Minimal Recursion Semantics*

Minimal Recursion Semantics (Copestake et al., 2005) is a description language for natural language semantics which is tailored towards the needs of an implemented grammar. MRS structures allow the underspecification of logical forms. This facilitates computational efficiency of syntactic processing with no loss of information because MRSs can be mapped to sets of fully-fledged logical forms by algorithm when applications require this resolution. Arguably, syntax only partially constrains sentence-level semantics in the majority of cases, so MRS provides a granularity of description which is appropriate for syntactic processing (Copestake et al., 2005).

3.5.1 *Overview and definitions*

Computational efficiency and underspecification is achieved by collecting hierarchical and relational information into flat lists.²² The information added to these lists provides constraints on the hierarchical and relational interpretations available. MRS underspecifies semantic information when more than one logical form is compatible with the constraints contained in the lists.

²²From the technical definition of MRS, these objects are unordered multisets or “bags”. However, lists are often used in implementations.

This is illustrated by example below in this section after the definitions.

There are two main lists in the MRS of Copestake et al. (2005). The first collects items called elementary predications (EPs; in some grammars, the list of EPs is called `RELS`). These are n -ary relations which are used to represent a (heterogeneous) collection of semantic notions. Some EPs stand in for lexical semantics. These lexical predicates can be logically unary. For example, the lexical meaning of *apple* might be represented via an EP called the “apple relation”. Such a unary EP would stand-in for a semantic predicate which is true of a semantic index x where x is an apple. Lexical EPs may be logically binary: a “build relation” can represent the semantics of a two-place logical predicate of x and y which is true when x and y are the two distinguished role players some particular building relation. EPs may have higher arity. EPs may alternatively correspond to abstract, grammatical predicates such as a “negation relation” — the relation logically inverts a truth value — or a “copula identity relation” — the relation which holds between two arguments which are grammatically indicated as identical (as in *the morning star is the evening star*).

Traditionally in DELPH-IN practice, the relation names of Elementary Predications are structured tags which provide lemma-like information about their lexical origin. The first row of Table 3.2 shows the template of the structure. The labels of lexical predicates have a leading underscore which sets them apart from grammatical relations. This is followed by a lemma, a part of speech tag, a sense number and finally, the suffix `_rel`. Abstract, grammatical predicate names do not bear the leading underscore.

The language (and notation) of MRS structures was influenced by the Hole Semantics of Bos (1995), which is also aimed at representing underspecified information. One of the key insights of Hole Semantics is that the creation of meta-variables or labels which can refer to formulas can be used to independently specify how semantic information would be tied together without having access to complete information. Thus elementary predications which appear in an MRS structure are also given a “handle”, a symbolic label which is used to refer to them elsewhere in an MRS.

Scopal argument positions of semantic functions take handles as arguments. These handle-

Wordform	Predicate label
“WORD”	(_)LEMMA_POS-TAG_SENENUM_rel
“apple”	_apple_n_1_rel
“あべこべ”	_abekobe_a_2_rel
“eat”	_eat_v_1_rel
“k ^w si”	_k ^w si_rem_q_rel
“n’t”	neg_rel
“is”	copula.id_rel
“d-”	_d_x_1sg.poss_rel

Table 3.2: Structured Labels in DELPH-IN Elementary Predications (<http://moin.delph-in.net/RmrsPos>)

typed arguments are referred to as “holes”. Scopal relations are expressed in terms of these holes and handles. The list of scope constraints, or handle constraints (HCONS) is contained in the second list. Handle constraints express scopal relationships which must obtain in a fully-resolved semantic tree. Because generalized quantifiers (Barwise and Cooper, 1981) are also treated as elementary predications which can appear in a RELS list, they also bear a handle to which scopal constraints may refer. The example in (16a) shows the structure used to represent generalized quantifiers as EPs. They specify a bound variable, as well as a restriction and body.

(16)
$$\left[\begin{array}{ll} \textit{quantifier-relation} & \\ \text{HANDLE} & \text{h0} \\ \text{PRED} & \textit{some_rel} \\ \text{BOUND-VARIABLE:} & \text{e1} \\ \text{RESTR} & \text{h2} \\ \text{BODY} & \text{h3} \end{array} \right]$$

More formally, MRSs are defined as follows (Copestake, 2002; Copestake et al., 2005, 12-14, 299, respectively). The global and local top handles are discussed further below.

1. an elementary predication (EP) has four parts:

- (a) a “handle” which functions as a label
- (b) a named relation
- (c) a list of zero or more semantic arguments
- (d) a list of zero or more scopal arguments

Elementary predications are often written as *handle : relation (args)*.

2. Within an MRS, the list of EPs represents an implicit conjunction
3. An MRS structure is a tuple $\langle GT, LT, R, C \rangle$ where
 - *GT* is the “global top handle” (a distinguished handle which is not outscoped by any other handle in resolved scope trees)
 - *LT* is the “local top handle” (a distinguished handle which is not the label of a floating EP, such as a quantifier)
 - *R* is the list of EPs
 - *C* is a list of handle constraints which are used in computing scope trees

Monotonically adding to collections of predications and scopal constraints is compatible with the generation of traditional scope trees, but avoids the necessity of keeping and maintaining all candidate scope trees during syntactic parsing. In this way, MRS provides a way for semantic composition in a syntactic grammar to keep track of only the information contributed, leaving anything else underspecified. Other contextual factors which may play a role in calculating specific model-theoretic meaning don’t have to be represented in a grammatical structure nor in the output of grammatical parsing. In later MRSs, a third list of individual constraints (ICONS, Song 2017) has been adopted by the DELPH-IN community to capture information structure properties which do interact with grammatical phenomena.

3.5.2 *Representation in Typed Feature Structures, usage in Typed Feature Structure grammars*

Encoding of Elementary Predications

MRS structures can be encoded into typed feature structures within a TDL grammar (24). The example in (16) above shows an elementary predication represented in AVM format. Within many DELPH-IN grammars,²³ EPs are defined in a sub type-hierarchy rooted in the type relation (17a), which has subtypes along an arity dimension as well as an event-individual contrast (17b).

(17)

- a. `relation := avm &`
 `[LBL handle,`
 `WLINK list,`
 `PRED predsor t].`

²³Specifically, those based on the Grammar Matrix (Bender et al., 2002, 2010). Chapter 4 introduces this cross-linguistic grammar prototyping platform and its usage in the *řabřucid* project.

```

b. arg0-relation := relation &
   [ ARG0 individual ].

arg1-relation := arg0-relation &
   [ ARG1 semarg ].

arg12-relation := arg1-relation &
   [ ARG2 semarg ].

arg123-relation := arg12-relation &
   [ ARG3 semarg ].

arg1234-relation := arg123-relation &
   [ ARG4 semarg ].

event-relation := arg0-relation &
   [ ARG0 event ].

arg1-ev-relation := arg1-relation & event-relation.
arg12-ev-relation := arg12-relation & arg12-relation.
arg123-ev-relation := arg123-relation & arg123-relation.
arg1234-ev-relation := arg1234-relation & arg1234-relation.

; Noun relations

noun-relation := arg0-relation &
   [ ARG0 ref-ind ].

; etc

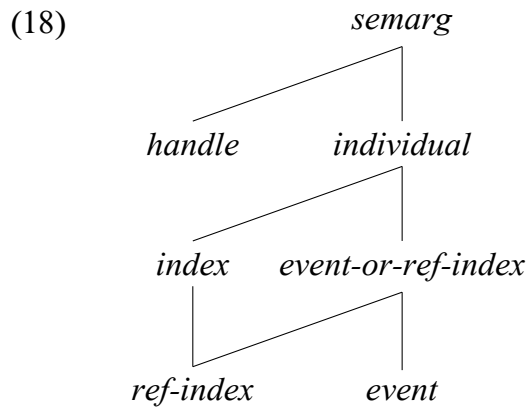
```

A further note is required with respect to the ARG0 properties illustrated in (17b): this property is interpreted via tradition as providing an index for the relation itself. Thus the specialization of the ARG0 property into ref-ind or event is concomitant with the types for noun-relation and event-relation, respectively. The values found in the ARG0s of EPS form the collection of values which can be assigned to an INDEX. The INDEX is one part of the HOOK structure: the collection of semantic information which a construction “publishes” for usage in combination. Discussion of these ARG0s values forms a central part of the case study in Chapter 7 and hook structures are presented below in this chapter.

Variable types and properties

In HPSG grammatical modeling, certain grammatical properties are usually associated with the semantic INDEX of a sign. The underlying motivation is that in particular languages, grammars encode a syntactic reflection of properties such as person, number, gender, etc. on wordforms and may impose particular syntactic constraints (usually related to agreement) on these values across a phrase. The first step in capturing these linguistic contrasts is to create a hierarchy of semantic types, or variables. Then, particular features may be assigned to these types with the grammar. While the organization of grammatical information is the standard *modus operandi* in HPSG analysis, this hierarchy of semantic types bears special status because they are mapped directly into MRS. Since MRS is really its own framework with its own ontology of concepts, these types represent a nexus of the syntactic theory of HPSG and theory which backs the semantic representation language of MRS. In this section I provide some background on these types in order to facilitate some of the detailed discussion in the case studies which are presented later in the document.

The hierarchy of types which can occur as values of features of EPs is presented in (18). The types appropriate for INDEX values are subsumed alongside the types appropriate for LABELS under the type *semarg*. Then, the hierarchy effectively provides a single further distinction of interest under the *individual* types into referential indexes (*ref-index*) and events (*event*). This is a basic distinction, in some sense, into predicates and (non-predicative) arguments. The *ref-index* type may index event-like phrases in nominalization or gerundial constructions. The *event* type may index statives or other predicates which are not truly events in the theory of *aktionsart*. The hierarchy, as is common in these types of models, provides intermediate types as well (*event-or-ref-index*, *index*) in order to allow partial information to be recorded.



In presentation formats of MRS, these type-names are traditionally mapped to single-letter variable names as shown in the mapping of (19). These abbreviations of variables types are then usually combined with a number to form a unique identifier for a variable. For example, an MRS presented with the variable names e13 and h3 is showing an event and handle type variable, respectively. This convention is offered by MRS processing software and throughout this document for convenience, but it does not form part of the formalism.

(19)

event	<>	e
ref-ind	<>	x
individual	<>	i
handle	<>	h
non_event	<>	p
*	>>	u
semarg	<<	u

By default, in Grammar Matrix derived grammars, the value for a *synsem*'s INDEX is constrained to *event-or-ref-index*. The variable-type hierarchy and its usage will come into play when discussing the predicate structure of Lushootseed verbs and valence-increasing morphology in the case study of Chapter 7.

Having introduced variable types, I can now turn to the idea of variable properties. The phenomenon of agreement is the canonical motivating example for providing the variable types with features. The features on variable types constitute a portion of the syntax-semantics interface in

that they record the semantic information about individuals, both eventualities and role-players, which a grammar makes use of in constraining agreement or other grammatical properties. Because these variables have a life outside of the syntax (i.e. they form what the grammar “exports” about the semantics) their features are referred to specially as properties.

Typical of *ref-ind* variables would be properties concerning person, number or gender; typical of *event* variables would be properties concerning tense, aspect or mood. The mapping of these properties from grammar-internal to grammar-external values is an ongoing topic of research in transfer-based machine translation (Bond et al., 2011). The motivation of a grammar-internal representation alongside a mapping to a grammar external representation is captured by phenomena like natural gender in contrast to grammatical gender: where the grammatical genders of noun classes constrain syntactic derivation language-internally, but do not provide cross-linguistic semantic information language-externally.²⁴ There are, of course, parallel examples concerning other properties of nominals (number, animacy, etc.) and properties of eventualities (tense, aspect, mood, etc). The categorization of variable properties for Lushootseed is of particular interest in the case study presented in Chapter 8.

Semantic Composition and Combination

The collection of EPs into monotonically growing lists,²⁵ is carried out via the *diff-list* structures (described above). Semantic composition requires more than just appending to lists, however. It’s not enough to model the sentence *Kim builds a house* as the append of the meanings of the individual components: [*Kim, builds, a, house*]. Functions and arguments must be combined so that *Kim* is the builder and *the house* is the built thing and other semantic information which

²⁴The topic of connecting real-world information about an entity to the grammatical properties of a linguistic objects is an especially interesting one. See the papers in Craig 1984 for an overview of the topic. See also Copestake (1992) on the ability of real-world information to structurally disambiguate, and the inapplicability of such a principle in grammar engineering.

²⁵Copestake et al. (2005) contains an endnote (8) which suggests that non-monotonic MRS is possible but not pursued. Further context is provided in Copestake et al. 2001, which encourages monotonically additive semantics because of interaction with generation algorithms. To wit, if semantic information can be dropped during derivation, generation algorithms which begin from the final semantic representation would not have an efficient search strategy in creating candidate intermediate forms.

is constrained by grammar (such as the cardinality of sets, time, aspect, modality) must also be encoded. In TDL grammars, semantic composition proceeds via grammatical rule-types which, upon unification with head and dependent lexical structures, co-index substructures found in dependent's `HOOK` into one (or more) of their own local semantic properties. For example: in *Kim sleeps*, the intransitive verb type of which *sleeps* is an instance will co-index its subject's `LOCAL|CONT|HOOK|INDEX` into the `ARG1` of the sleep relation that it provides. This co-indexation approach, of course, means that unification may rule out incompatible constraints on variable type and their properties.

Another type of information which must be encoded in the output of semantic composition is the relation of variable binding scope to the collection of predicates. The scopes of variable bindings may be thought of as a tree and predicates may then be seen as standing in conjunction on the nodes of that tree (as described in Copestake et al. 2005). In most MRS grammars implemented in TDL this idea is represented via label sharing (also as proposed in Copestake et al. (2005)). So restrictive modifiers will co-index their `HOOK|LTOP` to that of the item they modify, thereby sharing a node in a fully-resolved scope tree (and thereby forcing the conjunctive interpretation of their predicates within the same variable binding as that of their modificand).

It may be useful to see this in slightly more detail. The semantic information which a linguistic element presents for composition is encoded in a structure called *hook*. The definition of *hook* is shown in (20). The construction of a *hook* value for a particular grammar forms a hypothesis about the syntax-semantic interface for that grammar.²⁶ The hook provides the MRS top handles as well as the `INDEX` structure. The value of `INDEX` is, for lexical items, usually associated with a `KEY-RELATION` — a particular EP which provides the item's lexical semantics (Copestake et al., 2005, 299–300). The `ARG0` variable of this `KEY-RELATION` is typically co-indexed with the *hook*'s `INDEX`. value. Then, in composition, the `INDEX` is inherited up the head-path according to the Semantic Inheritance Principle of Sag et al. (2003). The `XARG` feature tracks a pointer into the `INDEX` value of the syntactic subject — this is in connection to raising phenomena and

²⁶See Copestake et al. 2001 for a mathematical description of this interface and it's relation to composition in the purely semantic domain.

constitutes a hypothesis that only the index of the subject, but not that of any other dependents may be accessed non-locally.

```
(20) hook := avm &
      [ GTOP handle,
        LTOP handle,
        INDEX individual,
        XARG individual ].
```

The types in (21) are typical of lexical entries. They map the lexical semantics from the `KEYREL` into in the local `HOOK`.²⁷ The `KEYREL`'s handle is identified with the `LTOP` feature of the `HOOK` structure and the `ARG0` is identified with the `INDEX`.

```
(21) norm-ltop-lex-item := lex-item &
      [ SYNSEM [ LOCAL.CONT [ HOOK [ LTOP #ltop ],
                                RELS.LIST.FIRST #keyrel ],
        LKEYS.KEYREL #keyrel & [ LBL #ltop ] ] ].

norm-hook-lex-item := norm-ltop-lex-item &
      [ SYNSEM [ LOCAL.CONT.HOOK.INDEX #index,
        LKEYS.KEYREL.ARG0 #index ] ].
```

Then, combinatoric lexical types provide a coindexation mapping from the `HOOK` features of their `ARG-ST` lists into their own `KEYREL` (or other) relations, as shown in the definition of *transitive-lex-item* (22).

```
(22) transitive-lex-item := basic-two-arg-no-hcons & basic-icons-lex-item &
      [ ARG-ST < [ LOCAL.CONT.HOOK [ INDEX ref-ind & #ind1 ] ],
                [ LOCAL.CONT.HOOK [ INDEX ref-ind & #ind2 ] ] >,
        SYNSEM [ LKEYS.KEYREL [ ARG1 #ind1, ARG2 #ind2 ] ] ].
```

So, under this system, lexical items expose, or publish, certain bits of semantic information in their `HOOK` feature. Items with unsaturated valence requirements implement the semantic part

²⁷ `KEYREL` values are also proposed for use in the grammatically relevant subset of semantic selection phenomena discussed in Flickinger and Bender (2003).

of combination by coindexing the INDEX of their dependents *hook* with the relevant ARG_N feature of their KEYREL (or, in principle, some other local relation).

Holes in handle-constraints are filled in a similar way. This is illustrated by the supertype intended for use by most determiner types (23). The QEQ handle constraint is added to the HCONS list and a quantifier relation is added specified as its KEYREL (note that matrix grammars use ARG₀ of quantifier relations to represent the bound variable of a quantifier relation). The quantifier's own RESTR handle is identified with the QEQ's HARG and the LARG is pulled from the head's HOOK's LTOP.

```
(23) basic-determiner-lex := norm-hook-lex-item &
      [ SYNSEM [ LOCAL [ CAT [ HEAD det,
                            VAL.SPEC.FIRST.LOCAL.CONT.HOOK
                            [ INDEX #ind,
                              LTOP #larg ] ],
                            CONT [ HCONS <! qeq &
                                   [ HARG #harg,
                                     LARG #larg ] !>,
                                   RELS <! relation !> ] ],
                            LKEYS.KEYREL quant-relation & [ ARG0 #ind,
                                                            RSTR #harg ] ] ] ].
```

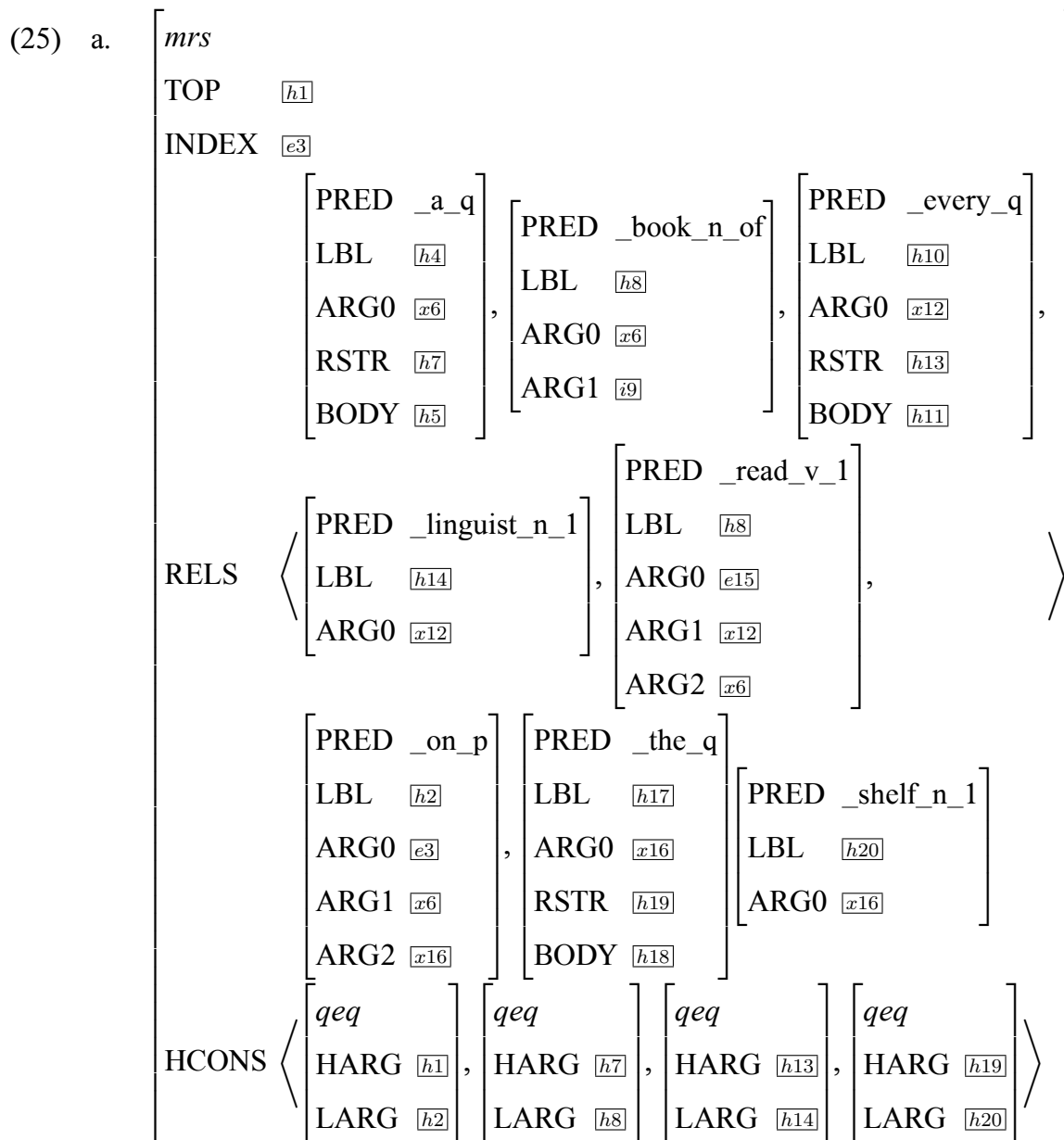
3.5.3 Putting it all together

The preceding outlines how semantic composition usually works in a DELPH-IN grammar. Above, I outlined the definition of an MRS as a four-tuple consisting of a global top handle, a local top handle, a bag of elementary predications and a bag of handle constraints (Copestake et al., 2005, 305). In this section I have presented two aspects of their usage in TDL grammars: the representation of EPs and the publication of ARG₀s as HOOK information which is co-indexed into local EP ARG_N features by lexical types with valence dependencies, but I have not yet shown the full representation of an MRS as a TDL structure. This is rectified in (24). This is the type for the value of the SYNSEM|LOCAL|CONT feature of most DELPH-IN grammars. The HOOK structure contains the top handles for the MRS, global and local, and the semantic index variable.

(24) `mrs-min := avm.`

```
mrs := mrs-min &  
  [ HOOK hook,  
    RELS diff-list,  
    HCONS diff-list,  
    ICONS diff-list ].
```

In (25a), a full MRS for the sentence “a book every linguist read is on the shelf” is presented in AVM format. Because it is often convenient to write MRS more compactly, an inline format is often used. This more compact format is shown in (25b). The two examples present equivalent information.



$$\text{b. } \left\langle \begin{array}{l} h_1, e_3, \\ h_4: \text{a_q}(\text{ARG0 } x_6, \text{RSTR } h_7, \text{BODY } h_h), \\ h_8: \text{_book_n_of}(\text{ARG0 } x_6, \text{ARG1 } x_9) \\ h_{10}: \text{_every_q}(\text{ARG0 } x_{12}, \text{RSTR } h_{13}, \text{BODY } h_{11}), \\ h_{14}: \text{_linguist_n_1}(\text{ARG0 } x_{12}), \\ h_8: \text{_read_v_1}(\text{ARG0 } e_{15}, \text{ARG1 } x_{12}, \text{ARG2 } x_6), \\ h_2: \text{_on_p}(\text{ARG0 } e_3, \text{ARG1 } x_6, \text{ARG2 } x_{16}), \\ h_{17}: \text{_the_q}(\text{ARG0 } e_{16}, \text{RSTR } h_{19}, \text{BODY } h_{18}), \\ h_{20}: \text{_shelf_n_1}(\text{ARG0 } x_{16}) \\ \{ h_1 =_q h_2, h_7 =_q h_8, h_{13} =_q h_{14}, h_{19} =_q h_{20} \} \end{array} \right\rangle$$

3.5.4 Conclusion

This goal of this section has been to introduce the basic principals and representation of MRS in DELPH-IN grammars. §7.2.2 of Chapter 7 continues the exposition of MRS when it explores some of the specific flavors of MRS which have been developed for particular use cases such as robust parsing with shallow tools or integration with dependency semantics.

3.6 Conclusion

This chapter has introduced a number of topics and representations which underly the grammatical presentations in later chapters. Starting from the principles of Formal Language Theory I have quickly progressed to a bifurcated model with Finite State Grammar for morphophonology and HPSG for morphosyntax and syntactico-semantics. Each of these frameworks, and their corresponding notations has been briefly introduced. These frameworks provide the guiding principles from a high level, but in order to understand the grammar development work which is described in later chapters, I must also introduce the tools and development environment which was used in building the project, which is the topic of the next chapter.

Chapter 4

ENVIRONMENT AND DEVELOPMENT CYCLE

In the preceding chapter, I described the theoretical frameworks and theories which underlie the work described in this dissertation along with a starter description of the tools and description languages at play. This chapter dives deeper on the latter topics, focusing on the tools and implementations used, the environment and the development cycle, and the overall architecture of the system.

4.1 Overall Architecture

The preceding chapter referenced the architecture of Bender and Good (2005). Bender and Good consider whether, in an implemented system, morphophonology should be “nested within” a morphosyntactic grammar, or whether morphosyntax should be “nested within a morphophonological grammar”. The former approach is relatively standard in HPSG — many grammars contain a `PHON` feature which collects phonological representations. In the LKB platform (Copestake, 2002), technical machinery is provided to do certain types of phonological transformations. The fact that any phonological transformations are written on types defined for morphosyntactic purposes means that morphophonology is effectively nested under morphosyntax. The opposite approach is seen in illustrative material for Finite State Grammar (Beesley and Karttunen, 2003; Hulden, 2009) where a finite state transducer adds tags indicating morphosyntactic word-classes underlying forms as they are analyzed. The difficulty with either of these approaches is that in natural languages, the phenomena to be described are most naturally and descriptively expressed independently. That is, in general, morphophonological rules tend to cross-cut morphosyntactic classes and vice-versa. For that reason, Bender and Good (2005) suggest an architecture with a shared lexical resource which is used to hold the information which is common to both systems:

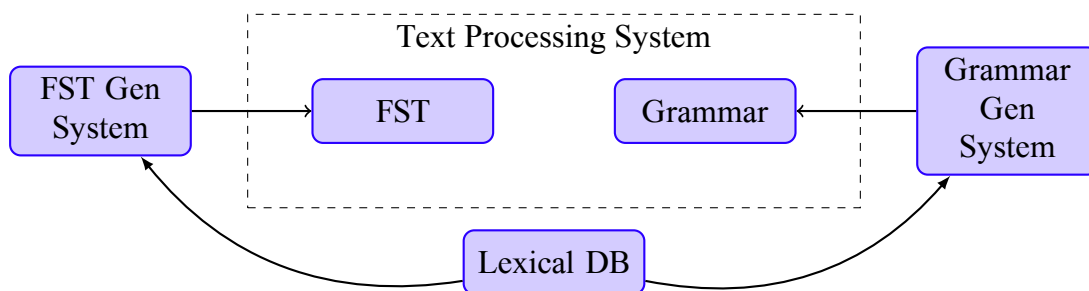


Figure 4.1: The System Architecture: a shared lexical database and two meta-grammatical systems to build a text-processing system.

the lexical items in their underlying forms, along with any idiosyncratic information about which classes they belong to. This project follows that approach — I have built independent systems for morphophonological analysis and morphosyntax which can, in principal, share a source lexical database.

For reasons articulated below, neither the finite state transducer which analyzes morphophonological rules nor the HPSG grammar used for morphosyntax are constructed directly, but are emitted automatically via a metagrammar generation system. The overall architecture of the shared lexical database providing input to both grammar generation systems and their output is presented pictorially in Figure 4.1.

The text processing system shown at the center of Figure 4.1 forms a bi-directional pipeline connecting on one extreme Lushootseed sentences in the language’s standard orthography to semantic representations at the other. Figure 4.2 illustrates this detail. The interface representation is the lattice of canonical, or morphophonemic forms.

In practice, I found that a bit of preprocessing was necessary in order to prepare flowing Lushootseed text for analysis. In the case of building the FST for morphophonology, I wrote a series of Python¹ and Bash² scripts to segment and prepare Lushootseed texts which were encoded in UTF-8 text files. In the case of testing and building out the HPSG grammar, test suite

¹<https://www.python.org/>

²<https://www.gnu.org/software/bash/>

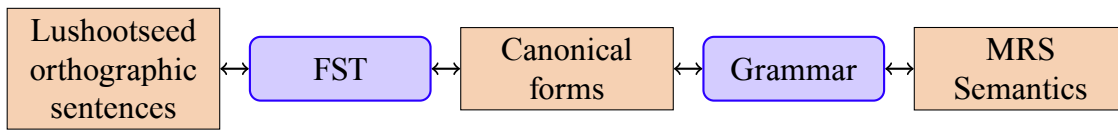


Figure 4.2: The text processing system as a pipeline.

profiles were prepared (again, by script). These scripts aren't directly represented in the figures showing the system architecture, mainly because that work is more of a practical nature than theoretical. More is said below on these profiles themselves and their usage in the development cycle (§4.7).

4.2 *The LinGO Grammar Matrix*

A derivative of the LinGO Grammar Matrix (Bender et al., 2002, 2010) is the central tool used by this project to generate implemented grammars. More specifically, the grammar generation system I use is a fork of the Grammar Matrix and customization system which has been specialized for Lushootseed grammar generation. The Grammar Matrix consists of a core-grammar which provides a central hierarchy of types which are posited to universally, or nearly universally applicable alongside a set of libraries which provides customization routines to add specific types which implement language-specific phenomena in a typologically motivated way.

The Grammar Matrix core grammar rolls up and packages a history of implemented grammars which precede it. That is, the main type hierarchy provided by the Grammar Matrix core grammar is a distillation of the type hierarchy implemented in the English Resource Grammar (Flickinger, 2000) with reference to the Jacy Grammar (Siegel and Bender, 2002) of Japanese. Bender and colleagues examined these preceding, large-scale, grammar projects and looked for types which could be feasibly useful in any grammar of natural language. For example, type definitions which implement monotonic semantic composition up a phrase structure tree (providing, essentially, the Semantic Compositionality Principle of HPSG (Sag et al., 2003)), or types which identify the `HEAD` value of a headed phrase with the `HEAD` value of the head-daughter (implement-

ing, essentially, the Head-Feature Principle of HPSG) are arguably useful across languages and are therefore provided by the main type-hierarchy of the Grammar Matrix.

One of the functions of the Grammar Matrix is to provide a resource for building starter HPSG grammars. The motivation for such a resource is built on the observation that:

1. Even languages which are typologically very divergent have huge overlap in terms of the grammar code necessary to implement them: things like the head-feature principal, semantic compositionality, or the usage of valence lists to model argument dependencies are hypothesized to be universal;
2. For languages which are typologically similar, even more structures and implementation may be shared;
3. Typological diversity itself is modelable and often parametrizable. For example, collections of languages may be grouped by asking whether they mark yes-no questions with a particle. Languages in the yes column may then be subgrouped, according to whether this particle is placed before or after the sentence, or somewhere else;
4. Grammars which share a common core are comparable.

These observations led to the building of a grammar customization system (Drellishak and Bender, 2005) which is made up of a core grammar file containing types which are hypothesized to be universally useful or applicable, a set of libraries which provide parameterized output for particular phenomena for a range of typologically documented parameters, and a questionnaire — deployed as a Web-application³ — which allows a user-linguist to provide the parameters to the libraries and then download a starter grammar package, ready to further customize and extend. The drawing in Figure 4.2 shows how the customization system works in some detail. A user-linguist answers a series of questions about the phenomena exhibited by the language

³<http://www.delph-in.net/matrix/customize/matrix.cgi>

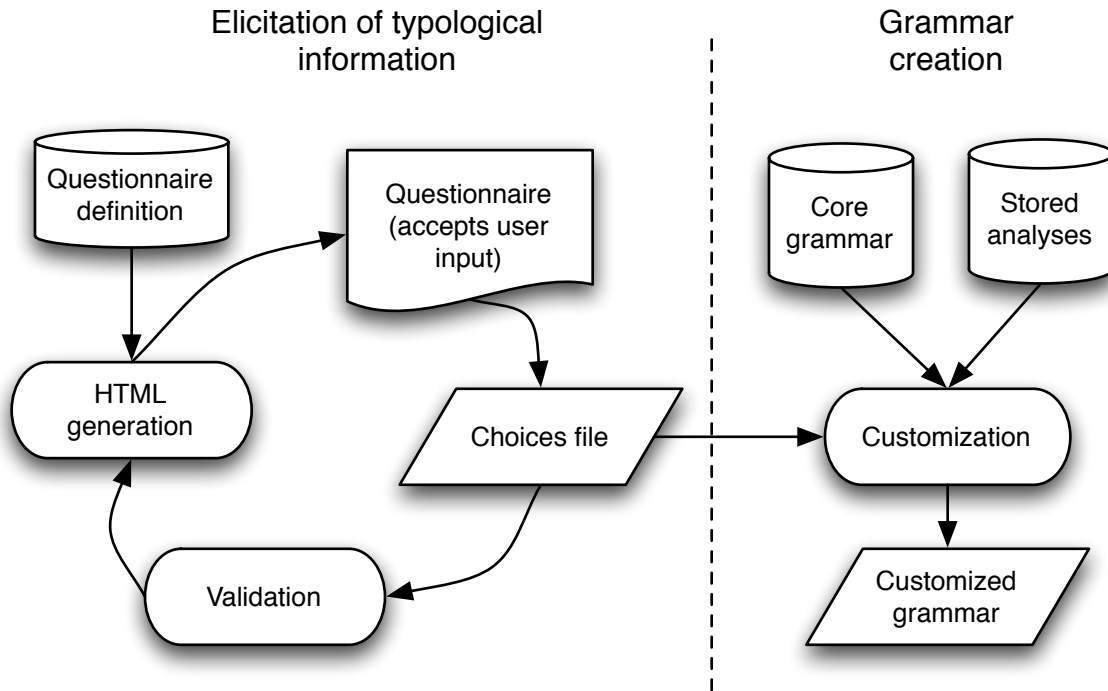


Figure 4.3: The customization of a grammar — figure from Bender et al. (2010)

they wish to model. The user’s answers to the questionnaire are stored as a key-value pairs in a format referred to as a choices file. Because there are interdependencies among choices which must be fulfilled in order to build a working grammar, choices files created on the web are subject to a validation system which provides feedback about fixing any issues. A choices file which passes validation can serve as input to the customization scripts. These scripts assemble types and instances computed from the choices along with the core grammar and package everything up for download. The packaging includes the configuration files necessary to load the grammar into one of the DELPH-IN compatible grammar processing systems.

In addition to the above merits, the Grammar Matrix customization system is free software, provided under the MIT license,⁴ so it can be used, distributed, and modified by anyone for any purpose. As mentioned above, the Lushootseed grammar generation system presented in this

⁴<http://www.delph-in.net/matrix/LICENSE.txt>

thesis is a fork of the Grammar Matrix customization system which is specialized in particular ways for generating a Lushootseed grammar. The open-source license allows this sort of derivative work.

In the original scenario imagined by the authors of the Grammar Matrix customization system, after taking advantage of the running start provided by the system, a user-linguist would continue their grammar development by directly editing the TDL. However, as discussed in the next section, there are advantages to remaining within a grammar generation paradigm.

4.3 Metagrammar Engineering: the CLIMB approach

Fokkens (2011); Fokkens et al. (2012); Fokkens (2014) lay out the motivations and basic methodology for building a metagrammar system using the Grammar Matrix customization system as a base. Fokkens named this methodology CLIMB (Comparative Libraries of implementations with Matrix Basis).

The crux of the idea is that grammar engineering is subject to path dependence.⁵ Design decisions which are taken early have an inordinate influence on the options available for later decisions. The older a decision is, the more “baked-in” it becomes to the fabric of the grammar; the harder it is to unwind. The process of grammar engineering is usually based upon a cycle of (1) examination of data, (2) identification of structural phenomena within that data to be analyzed, (3) proposing and testing analyses, (4) and then a return to (1). When proceeding along these lines, during any particular iteration a grammar engineer may discover a phenomenon which would have bolstered support for a previously unchosen analysis, had this data been under examination at that time. Because grammar engineers cannot look at all the relevant data and all the possible phenomena simultaneously, their work suffers from path dependence.

Meta-grammars alleviate this by flattening the decision tree into a single piece of editable code. Traditionally, a grammar engineer deciding between analysis A and B for a particular phenomenon would consult the available data and the current state of the grammar, make a con-

⁵The term “path dependence” from economics (Puffert, 1999; David, 1985), is also discussed in other fields such as history, law, another topics in the humanities (Margolis and Liebowitz, 2000).

crete decision about which analysis works best and then move forward to another phenomenon. In contrast, under the CLIMB methodology, the engineer does not make a hard decision, but instead builds both analyses A and B into the metagrammar system, and adds a parameter to the generation program to choose between them. This means that if, at a future date, the grammar engineer wants to unwind their choice and go with the alternative, the unwinding is done by setting new parameter choices in the metagrammar system, and rebuilding the grammar.⁶

For engineers working within the framework provided by the Grammar Matrix customization system, this means that instead of downloading their starter grammar and then going off to working directly on the grammar TDL, they instead remain within the framework of a grammar generation system, adding analyses and choices to their fork of the customization system.

In my case, this meant that I modified the word-order section of the Grammar Matrix to accommodate the particular variant of VSO word order exhibited by Lushootseed (discussed in Chapter 6). I added new scripts to the customization system which provided types that accommodated valence-changing affixes (discussed in Chapter 7). I built extensions to the choices file processing system to allow the importation of external lexicon files (presented in the next section of this chapter). At the end of the day, my metagrammar system accepts choices files and outputs different versions of the Lushootseed HPSG grammar according to the parameters specified within the input.

4.4 *Meta-choices*

Even though the CLIMB methodology emphasizes grammar generation through code, static inputs in the form of choices files are still used. The typical interaction between these components in a CLIMB grammar, as I have implemented it, is that the grammar generation code provides

⁶This claim merits the caveat that the competing analyses must, to some degree, be maintained independently. Because maintaining multiple parallel analyses is expensive, eventually the grammar engineer will opt for a particular choice. Nevertheless, the CLIMB metagrammar engineering framework allow the grammar engineer to delay the decision point until they are comfortable that things are coalescing in a particular direction. This creates a great improvement over the alternative, in which each decision is made one at a time, and moving back to undo a particular decision necessarily entails redoing all of the work between the current state and when that decision was taken.

a new option which is configurable via the choices file. So, the manipulation of choices files becomes an important part of grammar development. In order to facilitate the management of choices files, I implemented a new option to allow the inclusion of choices from one file into another.

In more detail, the motivation for this feature is that as I undertook the writing of new grammar generation code to support new choices, I needed to test these new choices in various combinations with existing ones. In this sort of scenario, a handful of items may be varying, but the majority of the choices needed to create a grammar remain unaffected in the configurations at play (e.g. the lexical entries). Copying entire choices files for each of these variants is inconvenient and poor engineering practice.

For example, a typical Grammar Matrix style choices file is divided into sections which organize the information into categories such as argument optionality, case, tense-aspect-mood, negation and more. In order to experiment with a new choice regarding determiners, I might make a copy of a file containing the choices from each of these sections with a new determiner choice and name the file something like `new-determiners.choices`. But if later, I find that I prefer to take a different approach to the modeling of tense, each of the choices files which I have accumulated must be updated accordingly — a task which is tedious and error prone. A better approach is to store choices in a more granular format, and allow inclusions from other choices files. With choices file include statements, choices file which represents an experimental approach to determiners can be saved as `new-determiners.choices`. Then, instead of copying over values which may become out-of-date as the grammar improves in other dimensions of analysis, the new choices files can contain include statements to simply import them. This helps grammar developers manage which interacting phenomena they wish to test without making many superfluous copies.

The inclusion of choices files is especially useful because of the fact that in typically, the grammar's lexicon itself is input to the customization system through the choices file. As opposed to other choices-file sections which typically only contain a handful of choices, the lexicon and morphology sections contain a large number of entries. And in my experience these sec-

tions of the choices file grow very rapidly, as lexical coverage is required to scale the developing grammar to meet new examples. After implementing meta-choices in my fork of the grammar matrix, the main choices file for most of the variations looks relatively skeletal (26) — which means that the interesting parts have been parameterized in a way that maximizes code-reuse and flexibility.

```
(26) version=28

    section=general
    language=LushootseedMRS
    iso-code=lut-mrs pass-lkeys=true
    semantics=mrs
    determiners=nominalizing
    @import=punc-chars.choices

    @import=word-order.choices

    @import=png.choices

    @import=case.choices

    @import=arg-opt.choices

    @import=tam.choices

    @import=other-features.choices

    @import=coordination.choices

    section=lexicon
    @import=det-lexicon-NL.choices
    @import=det-lexicon-shared.choices
    @import=noun-lexicon.choices
    @import=verb-lexicon.choices
    @import=other-lexicon.choices

    @import=/morphology.choices
```

4.5 *Transducer Generation*

The lessons of the CLIMB methodology for minimizing path-dependence don't apply only to the building of syntactic grammars. The use of a shared lexical resource for both grammar systems implies that the Finite State Transducer (FST) ought to be able to be rebuilt by script whenever the lexical resource is extended.

More details about the development methodology employed to build out the FST are given Chapter 5. Here I want to briefly present the data format used to facilitate automatic generation of the FST from a lexical database. Recall from the previous chapter that a typical finite state transducer for morphophonological analysis is built from a concatenation of lexical resource files which model position class morphotactics to yield a network which is a finite state automaton. Then, because the automaton is also an identity transducer, this network can be then composed with a series of transducers which model morphophonological rules (typically, this network is also minimized for efficiency (Hopcroft et al., 2006)). So, as a starting point, the generation of an FST from source files will focus on providing the lexical items and information about their organization into position classes. The generation code can then export the resource in the format expected by the FST platform.

In order to make a start on a database format to support FST generation (as well as, eventually, morphosyntactic choices file generation), I enumerated the following properties about a morpheme which ought to be stored:

1. Morphophonological information:
 - (a) position class
 - (b) morphological class
 - (c) cophonology
 - (d) suppletive forms
 - (e) idiosyncratic affix selection

2. Morphosyntactic information:

- (a) syntactic type, including valence properties
- (b) predicate symbol
- (c) syntactic irregularities

This template is general, and represents an ambitious goal of what to record about an item. I began with a simple, home-rolled XML format for the lexical database. Using a plain text format like XML provides transparency of representation and relative ease of editing or porting to new platforms compared to the binary formats of key-value databases such as `dbm`⁷ or BerkeleyDB⁸ or relational databases.⁹

In fact, beyond the list of linguistic information mentioned above, to make an actual working export script which builds the FST from the database files, I need to provide particular information about the file structures, lexicon names, and other sorts of implementation details which are required to stitch-together a working system. So, the `fst-files` element contains information about the files to be generated and the lexica they contain, the `lex-items` element lists items and their properties.

An example of the XML utilized for the shared database is presented in (27).

⁷<https://www.gnu.org.ua/software/gdbm/>

⁸<https://www.oracle.com/database/berkeley-db/>

⁹While there are binary formats, nowadays, which are open and stable, and unlikely to be forgotten, the cautionary tales and resultant best-practices described in Bird and Simons (2003) are still relevant.

```
(27) <lexdb>
<fst-files>
  <file name="nmlz.lexc" id="nmlz" multichar-symbols="@P.NMLZ_V.ON@">
    <lexicon name="Root" hasnull="true"/>
  </file>
  <file name="asp.lexc" id="asp" multichar-symbols="@P.NMLZ_V.ON@">
    <lexicon name="Root" hasnull="true">
      <lexicon name="Rest" hasnull="false"/>
    </lexicon>
  </file>
</fst-files>

<lex-items>
  <item id="i0">
    <mp>
      <filename>nmlz.lexc</filename>
      <lexiconname>ROOT</lexiconname>
      <!-- ut: upper tape, lt: lower tape -->
      <ut flags="@P.IL.ON@ @0.OTHERFLAG@">XX</ut>
      <lt flags="@P.IL.ON@ @0.OTHERFLAG@" radical="n:m" >XX</lt>
      <!-- lower tape contains radical designation in some cases, just
use string splicing indicies (like Xigt, perl, python, etc) -->
      <!-- output gloss into a comment in generated files -->
    </mp>
    <ms>
      <!-- more to be added later, certainly -->
      <pred>_foo_n_1_rel</pred>
    </ms>
    <gloss>foo</gloss>
  </item>
</lex-items>
</lexdb>
```

The script which generates the lexicon files from this database is included in Appendix A. After the lexicon files have been generated from the source, another shell script calls the foma processor to build and optimize the binary executable.¹⁰

¹⁰Foma (Hulden, 2009; Hulden and Alegria, 2010) is a free and open-source reimplementation of the proprietary tools XFST and lexc described in (Beesley and Karttunen, 2003) which is, for the most part, is API compatible with those tools.

4.6 Testsuites, Grammar profiles, Regression Testing

Just as quality assurance in complex software is driven by tests which ensure the behavior of the system, best practice in grammar engineering is to reify implemented phenomena into batteries of regression tests (Oepen et al., 2002). These allow a grammar developer to confidently extend the grammar in ways which interact with previously implemented analyses. These regression tests usually take the form of testsuites which pair surface-form sentences with their (respective) expected output(s) after processing. When possible, testsuites include “negative”, or “ungrammatical” examples which the grammar is expected not to parse (these items have no expected output), as well as “positive” ones, for which the grammar is expected to return a particular semantics. Negative examples can be harder to come by than attested ones, since generating them almost always requires the validation of a native speaker. My own testsuites contain very few negative examples for this reason.

To facilitate test driven development within Grammar Engineering, the software package `[incr tsdb()]`¹¹ provides functionality to track the performance of a grammar on a battery of tests over time (Oepen and Flickinger, 1998). The letters “TSDB” in the package name make reference to an earlier program which provided a “testsuite database” for developing a grammar (Oepen et al., 1998) and the backend of the improved `[incr tsdb()]` package still requires that input sentences are stored in a plain-text database format which records information about them. `[incr tsdb()]` is a general purpose tool for grammar engineers. The schema for this database (see Table 4.1) is rich, providing the user a lot of opportunity to enrich examples with metadata. However, not all of these columns are required in order to operate the software. The testsuites constructed for this project make use of the input field, which is provided to an attached instance of a grammar processor upon testing, the wf (well formed) field, which records whether an example is positive or negative, and the author, date and comment fields, which provide locations to store provenience information.

In the context of Grammar Matrix-based development, a command line system has been

¹¹Pronounced [ti.es.di.bi.'plʌs.plʌs] (<http://moin.delph-in.net/ItsdbTop>)

id	integer (key)
origin	string
register	string
format	string
difficulty	integer
category	string
input	string
tokens	string
gloss	string
translation	string
wf	integer
length	integer
comment	string
author	string
date	date

Table 4.1: [`incr tsdb()`] fields and the corresponding data types available for each testsuite item

developed in order to allow multiple developers to develop libraries while ensuring backwards compatibility with previous library implementations (Bender et al., 2007). Because my system is not just a stand-alone grammar but is a fork of the Grammar Matrix customization system, I extended the regression testing system which is already built into the project in order to run Lushootseed related tests. This allows me to ensure that multiple, competing approaches to Lushootseed grammar continue to work even as I edit and extend the libraries that generate those grammars.

The built-in system stores each test as a triple: a choices file, a testsuite and the gold-standard semantics associated with that testsuite. Note that the built-in system does not test the resulting syntactic parses against a gold standard. The assumption here is that the semantic results are the valuable output of the system. The syntactic parse just serves as a means to an end (and therefore, can be ignored in regression testing).

One way in which my extended regression testing system differs from the built-in system is that the built-in system relies in several ways on the language name of the choices file for book-

keeping duties such as temporary directory names and pairing up choices files with testsuites. However, because I will be testing the same choices file against several testsuites, the system had to be modified to use a different identifier for a particular running test instance. Therefore, I forked the existing regression testing system into a new subdirectory and modified the testing scripts to my purposes. The result was that after creating new functionality in the metagrammar system, I can commit that functionality into a regression test and add it to the battery of tests to be run after each development cycle.

4.7 *Lushootseed data, testsuite construction*

The Lushootseed data used for building and testing this grammar project comes from the following sources:

1. The examples sentences from the first eight lessons (and the summary lesson) *Lushootseed Reader volume 1* (Hess, 1995, LR1)
2. An electronic version of the Lushootseed Stories published in Beck and Hess (2014, 2015)
3. Examples I constructed to fill out testsuites based on the prose descriptions of Lushootseed in the literature. These examples were generally provided to language experts David Beck and Zalmi *ʔaswəli* Zahir for validation.

The first part of LR1 is a pedagogical grammar which presents Lushootseed grammatical phenomena in a systematic way. For this reason, the examples provided often illustrate paradigms and contrasts. Most of the time, they are also short and simple, so that the student can focus on the phenomenon under discussion. This makes them ideal sentences for a grammatical test-suite. The example sentences from LR1 were transcribed from the book into a plain text file with the example sentence and its gloss appearing as pairs of lines. The total number of sentences transcribed was 548.

Because the examples in LR1 are presented in standard Lushootseed orthography, I also segmented the first sentences from the first ten lessons of the reader into morphophonemic forms

(including the unnumbered intermezzo which falls between Lessons Eight and Nine: “Summary of Function Marking”. These were then made into TSDB profiles and were eventually turned into the basis of regression tests for the HPSG grammar.

The Lushootseed Corpus in Beck and Hess (2014, 2015) contains a detailed interlinearization of traditional stories told by native speakers to Thom Hess in the middle of the 20th century. The format is a four line IGT in which the first line is standard Lushootseed orthography, the second line is morpheme segmented and morphophonologically regularized, the third line provides a linguistic gloss, and the fourth line is a free translation.

David Beck provided me a copy of this data in Rich Text Format,¹² encoded in UTF-8 with a Byte Order Marker. In order to prepare the data for use, I normalized punctuation (converting Unicode en-dashes and quotation marks to their basic ASCII counterparts), and then output the conversion in two formats: a language-only format which included only the original orthography, one sentence per line, and a plain-text IGT format which was essentially a normalized version of the originals. The former was used for generating input to the FST during testing, and was exported into the `[incr tsdb()]` database format to be used to create grammar profiles for the syntactic grammar development.

Finally, in order to handle the testing of an implementation of reduplicative morphemes, I created a new set of testsuites. This is because the original IGT from David Beck keeps the phonetic material from reduplication even in the morpheme regularized line — see the second line of (28). But at the level of morphosyntactic analysis, the form of reduplicative morphemes must have already been regularized to something generic, along the lines of what appears in the gloss line of Beck and Hess’ IGT (the fourth line of (28)). So, I wrote a script to generate a representation such as the one shown in the third line of (28). This is the format that my syntactic grammar expects. Morphemes appear in their underlying representations, which, in the case of reduplicative materials, has to be something like a gloss. I used a script to generate these such forms and then created testsuites for each of the four stories by Edward Sam which are included

¹²<http://www.microsoft.com/en-us/download/details.aspx?id=10725>

in LR1.

- (28) *tuχ^wax^w čəd ɬuyəhubtubicid ti tusyəhub ?ə turdi?*
tuχ^w=ax^w čəd ɬu=yəhub-tx^w-bucid ti tu=s=yəhub ?ə tudi?
tuχ^w=ax^w čəd ɬu=yəhub-tx^w-bucid ti tu=s=yəhub ?ə tudi?
 just=now 1SG.SBJ IRR=recite-ECS-2SG.OBJ DET PST=NMLZ=recite PR DET
tusluχ^lluχ^l čət
tu=s=luχ^l-luχ^l čət
tu=s=dstr-luχ^l čət
 PST=NMLZ=DSTR-old 1PL.POSS

‘I am simply going to recite to you a traditional story from our distant Elders.’

Mink and Tutyika (ES)

The list of testsuite profiles used in the development of the grammar, along with some descriptive statistics, is provided in Table (4.2).

4.8 Conclusion

This chapter has provided an introduction both to the overall architecture of the project — what the components are and how they should interact — and also the development cycle and tools used in the day-to-day work. This is the end of the introductory material which is intended to provide background. In the next chapter, I overview the construction of an initial finite state transducer for morphophonological analysis. In subsequent chapters, I turn to the development of the syntactic grammar.

Name	Items	Words-per-item	Source
LessonOne: -tx ^w , agent, patient	12	2.0	LR1
LessonTwo: čəd, čəx ^w , čəł, čələp	14	3.14	LR1
LessonThree: -d and -b, ti and tsi	19	3.21	LR1
LessonFour: -s and -d, goal; experiencer	36	3.25	LR1
LessonFive: -dx ^w ; summary of patient oriented suffixes	20	3.55	LR1
LessonSix: Patient suffixes + b, ?ə + agent	37	5.1	LR1
LessonSeven: Agent oriented stems, ?ə + patient	48	5.15	LR1
LessonEight: -yi- role, ?ə + recipient	24	5.13	LR1
SummaryOfFunctionMarking	27	3.7	LR1
LessonNine: person suffixes: including reflexive and reciprocal	52	1.98	LR1
LessonTen: stative perfective prefixes; verbs expressing cognition	16	3.125	LR1
-yi- role, ?ə + recipient	16	5.13	LR1
Mink and Tutyika	52	4.88	Edward Sam
Bear and Ant	93	3.69	Edward Sam
Coyote and Big Rock	80	4.33	Edward Sam
Bear and Fishhawk	97	3.9	Edward Sam
Subject Auxiliaries	56	3.5	mixed
Transitivizers	22	2.6	mixed
Object Markers	56	2	mixed

Table 4.2: Testsuites used in HPSG Grammar Development

- c. *g^wəl tastaʔlil tiʔiʔ kikəwič*
g^wəl tu=ʔas-ʔaʔlil tiʔiʔ ki-kəwič
 SCONJ PST=STAT-live DIST ATTN-hunchback

‘And Little Hunchback dwelled there.’

Basket Ogress (DM)

From the point of view of syntactico-semantic analysis, the alternations in (29) are irrelevant. The syntactico-semantic component only needs to know that the *STAT*- prefix is attached. Encoding the contextual spelling changes into the syntactic grammar would not only complicate the analysis, it would also make the grammar more difficult to edit and maintain (Bender and Good, 2005).

5.1 Resources, Methodology, Initial Results

My primary descriptive resources for this implementation were the pedagogical grammars by Hess and Hilbert (1995a,b), the grammatical snapshots in the first parts of Hess (1995, 1998) and the sketch of morphotactic position classes and selectional allomorphy in Beck (n.d.). The next step in resource acquisition was to acquire a lexicographic resource, preferably in machine-readable format, which would provide a listing of the roots, stems, affixes and clitics which populate the continuation classes of the morphotactics. Unfortunately, however, the principle lexicographical source on Lushootseed, Bates et al. (1994), is currently only available in paper form.² Thus while one can read about Lushootseed morphophonology in the literature and review the lexical items in paper form, the challenge of building up a reasonable-sized collection of machine-readable roots and stems in order to populate the underlying lexical files is a nontrivial one.

²Although Bates and Lonsdale (2010) describe the conversion of the dictionary from legacy software to a modern XML representation and Lonsdale and Matsushita (2011) make use of this electronic resource to build a two level model of Lushootseed morphophonology, neither the electronic dictionary nor the two-level model are available to the public or to other researchers at the time of this writing.

5.1.1 *Iterative Development*

Thanks to a collaboration with David Beck, I was able to answer this challenge by starting with an electronic version of the texts which were eventually published as Beck and Hess (2014, 2015). This collection of four-line interlinear texts provides both an entry point for building a lexical resource as well as a set of gold standard³ data against which the developing machine can be evaluated. Beck's digitized corpus can serve both these purposes because the second line of these texts presents a regularized, morphophonemic line, the input format which would be passed on to a syntactic grammar. Because this data was hand-created by experts of the language, it provides a standard against which the output of the automatic system being constructed can be evaluated. The methodology undertaken here, then, is summarized in Figure 5.1 and described in the following paragraphs.

The process began by selecting a story for preliminary development. "Basket Ogress" as told by *?alata?* Martin Sampson was chosen. The initial system was then set up by looking through the lexical items in the story and populating the lexicon files with roots and affixes as they are found in the text. As a starting point, an initial morphotactic system was adapted from the table in Beck (n.d., 30). The relationships between the lexical forms and orthographic forms were observed in the first text and phonological rules were posited and tested.⁴

Once the analyzer accepted 100% of the wordforms in the preliminary story and mapped them with 100% accuracy to the lexical forms in the gold data, evaluation upon unseen text began. A second story was chosen from the corpus, and the system as developed based on the data in the preliminary text was run against the second story. The following metrics were recorded: (1) coverage rates for types, tokens, (2) ambiguity, (3) accuracy against the gold standard. Coverage refers to whether or not the analyzer returns at least one analysis for a given form. Ambiguity, measured overall in readings per type, refers to the fact that there can be more

³In computational linguistics, the term "gold standard" is often used to refer to hand-created data against which the output of a computational system can be evaluated.

⁴Much other initial work was done at this stage to facilitate an overall development environment: scripts were written to do the updating of the system network, the running of the analyzer against texts, as well as evaluation of an output against the gold standard.

than one analysis for a given wordform. Accuracy against the gold standard refers to whether or not the output of the analyzer matches the analysis of that wordform in the second line of Beck and Hess (2014, 2015). In the case of ambiguity, the system's answer was considered correct as long as one of the analyses matched that of the gold standard. In order to measure the ability of the analyzer to generalize to unseen word forms, the metrics were also recorded for the analyzer after removing any word triples (form, lexical form, gloss) which are found in the preliminary texts.

Next, development continued against the second story: first, any wordforms which failed to be accepted because of missing stems in the lexicon files (either roots, affixes, lexical suffixes or clitics) were counted and those stems were added. Then, another set of measurements were taken. This measurement effectively extracts the error rate due to missing lexical material, leaving an evaluation of the system's combinatorics and transformations.

After this, the remaining forms which failed to be analyzed did so because of some morphotactic or phonological rule which was previously unimplemented. These rules were then added until coverage on the second story reached 100% with 100% accuracy as measured against the gold standard. At this point, the analyzer was ready to be tested against a third story. In principle, the steps of this iterative methodology are to be repeated continually against new texts until the coverage rates and accuracy begin to approach an upper limit. In fact, three complete cycles (plus the preliminary cycle) have been completed.

This methodology builds in a running evaluation which is intended to show the maturation of the system over time. By factoring out a measurement against only unseen forms, the evaluation shows how well the system generalizes. By measuring against new texts again after adding missing stems, I can look at how much of the error rate against a text is due to missing stems and how much is due to missing morphotactic or phonological rules.

In this way the expectation is that as the system matures, the error rates should asymptotically approach some low rate — not zero. The value for this rate is something that can be empirically established later on in system development.

0. preliminary:
 - (a) Choose an initial story for development
 - (b) Populate lexicon files with roots and affixes found in the text
 - (c) Set up morphotactic combinations based on preliminary table in Beck (n.d.)
 - (d) Define phonological rules to map lexical forms to surface forms based on the correspondences in the gold data
1. run morphological analyzer against a new story, record evaluation data
2. record evaluation data against unseen forms only
3. add missing stems, record evaluation data
4. modify morphotactics and phonology as needed
5. select a new text, return to step (1)

Figure 5.1: Enumeration of the FST development methodology

5.1.2 Initial Results

Having discussed the methodology and the evaluation strategy for this project, I can now present some initial results. The overview is that the preliminary step, and three full iterations of the evaluation cycle have been carried out. Further discussion follows the presentation of the numbers.

The preliminary text chosen was “Basket Ogress” as told by *ʔalataʔ* Martin Sampson (MS). This text has 83 sentences comprised of 395 word tokens which fall into 158 unique wordtypes. After the preliminary development stage the system captured all this data with 100% accuracy, finding 14 wordtypes to be ambiguous (average ambiguity per type of 1.12). That was the first iteration of the analyzer to be tested on unseen text.

“Basket Ogress” as told by Dewey Mitchell (DM) was selected for this test.⁵ DM’s telling consisted of 87 sentences, 428 wordtokens falling into 123 word types. The initial system applied to this text provided a word-type coverage rate of (36.27%). The rejected types consisted of 147 tokens, so only 34.35% of tokens were rejected (token coverage rate: 65.65%). That is, the types which were rejected were not high frequency types. The system found four ambiguous types with an average ambiguity of 1.04 per type. While many types were rejected by this initial analyzer, the types which did receive coverage were analyzed accurately: 98.93% of tokens which received coverage matched the gold data analysis.

The next step was to evaluate the performance of the system on unseen wordtypes. This test consisted of wordforms from DM’s telling after subtracting wordforms which occurred in MS’s telling, leaving 189 wordtokens falling into 159 wordtypes. As expected, the performance of the system on unseen words is necessarily lower than the performance of the system on the entire text. Coverage fell to 22.64% percent of types (22.22% of tokens). Two types were ambiguous

⁵The collection of texts in Beck and Hess (2014, 2015) contains six tellings of “Basket Ogress”, five tellings of “Star Child”, three tellings of “Mink and Tutyika”. I decided beforehand that during development, tellings of the same text would be added in sequence. Because of the measure on unseen wordforms only, this does not adulterate the view into the system’s ability to apply to generalize to unseen forms. That is, despite the narrative similarity, extracting the unseen forms shows that the tellings do not present exactly the same collection of words. See, for example, in Table 5.1 that ML’s telling of “Basket Ogress” contains 397 wordforms, 347 of which are not found in MS or DM’s telling.

giving an average ambiguity of 1.02 per type. While this coverage rate is rather low, even on unseen words, accuracy was still quite high at 92.86%. Although these initial numbers are low, they provide a baseline against which the system improve over time can be measured.

Next, the rejected wordforms from DM's telling were inventoried and categorized as to whether the failure was due to missing lexical material or missing morphotactic/phonological rules or both. Missing lexical material was added and the evaluation metrics were recorded again. Coverage rose to 73.58% of types (87.15% tokens) with ambiguity still relatively low at 9 types (1.07 average per type). Accuracy remained high at 98.39%.

After this, morphotactics and phonological rules were modified until analysis of the DM telling reached 100% coverage with 100% accuracy. This completed the first evaluation cycle.

For the second cycle, a third text was selected: Martha Lamont's (ML) telling of *Basket Ogress*. This text is significantly longer than first two with 240 sentences comprised of 987 word tokens distributed across 397 types. Coverage rose to 55.16% of types, which accounted for 74.77% of tokens. Ambiguity was noticeably higher at 1.15 readings per type. Accuracy remained high at 98.08%.

As before, the system was also measured after removing from ML's telling any wordform triples (orthography, lexical form, gloss) which occur in either MS or DM's telling. This left 342 types instantiated in 473 tokens. The new iteration of the system provided analyses for 48.15% of types (up from 22.64% against the unseen forms in the first iteration), or 46.57% of tokens (as compared to 22.22% in the first iteration). This is an encouraging result: the system's performance on completely unseen forms has improved by approximately 25%.

Finally, items which were rejected for missing lexical material were inventoried, and that material was added to the lexicon files and the tests were run again. This second iteration of testing after adding missing lexical material provided coverage for 87.15% of the stems in ML's telling of "Basket Ogress" (92.21% of the tokens). At this point, the system was further developed by to capture the remaining phenomena in ML's telling, completing the second cycle and the third cycle began.

In the final cycle completed to date, a fourth telling of "Basket Ogress", this one by Louise

Anderson, was selected. In this iteration coverage rose to 66.6% of the wordtypes (81.99% of tokens), and ambiguity held relatively stable at 1.15 analyses per type. Accuracy also remained high, at just over 98%.

Again, wordform triples which were found in any of the preceding stories were removed and the system was tested on unseen forms only. On this metric, the system remained flat from the previous cycle: coverage for types and tokens were nearly unchanged (exact measurements shown in Table 5.1 and Figure 5.2).

Again, the forms without analysis were inventoried and where a missing stem or lexical item was missing, these were added to the appropriate files. Then the system was tested against the entire text of LA's telling of "Basket Ogress" for the second time. Like the evaluation on unseen forms, the system's performance after adding new lexical material this cycle was nearly the same as last, although type coverage was slightly lower (and token coverage was slightly up).

The data in Table 5.1 shows the entire set of measurements to date; in order to better illustrate the trend, the chart in Figure 5.2 presents some of these data graphically. In the chart, the red indicates the overall performance of the system on unseen texts. The blue indicates the system's performance on only unseen words in an unseen text — this metric is intended to help show the system's ability to generalize. The green indicates the results of testing on the text after adding missing lexical material — intended to help show what part of the unanalyzed forms are related to missing lexical entries and what part is related to missing phonological or morphotactic rules.

5.1.3 Discussion

The preliminary results presented here are encouraging. They show a consistently high accuracy rate, due to the hand built nature of the system.⁶ Although coverage on the initial round taken against DM's telling was low, the second and third set of measurements taken against ML and LA's tellings shows us that things are moving in the right direction. Especially encouraging is

⁶Broadly speaking, in computational linguistics, much success has been achieved by constructing systems built on statistical models. There is a known complementary dichotomy whereby rule-based systems tend to be precise, often at the cost of coverage, and statistical models tend to provide high rates of coverage, often at the cost of precision.

	type coverage (%)	token coverage (%)	avg. ambig. (readings per type)	total types	accuracy (% types)
DM	36.27	65.65	1.04	193	98.93
DM.unseen	22.64	22.22	1.02	159	92.86
DM.stems	73.58	87.15	1.07	193	98.39
ML	55.16	74.77	1.15	397	98.08
ML.unseen	48.25	46.57	1.16	342	91.03
ML.stems	87.15	92.21	1.26	397	96.77
LA	66.66	81.99	1.15	198	98.36
LA.unseen	49.62	27.19	1.18	131	89.16
LA.stems	86.87	94.68	1.25	198	98.36

Table 5.1: The FST evaluation data to date

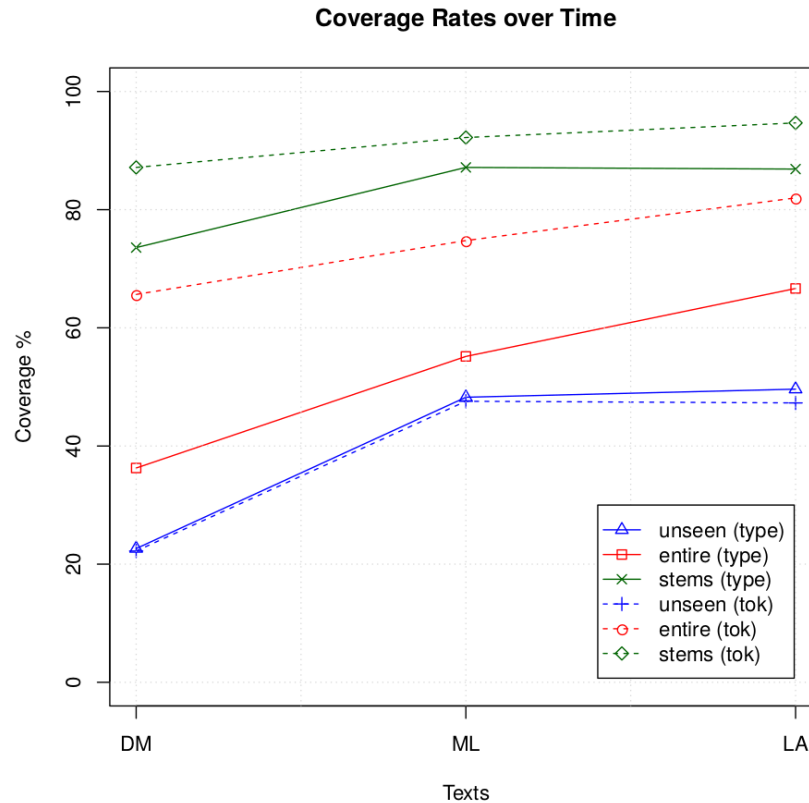


Figure 5.2: In this chart showing the initial performance of the system, solid lines show coverage rates against wordtypes, dotted lines against tokens. Red indicates performance on an entire text, blue indicates performance only on unseen wordforms, green shows the reevaluated performance on the entire text after adding new lexical material.

the range of improvement on the basic metric of coverage on entire unseen texts (from 36.27% to 66.6% coverage for types, with accuracy stable at 98%; indicated by the red line in Figure 5.2).

In this sort of evaluation scheme, each iteration of results can be viewed as a reference point against which later iterations can be compared. After further iterations of development and evaluation have been performed, the rate of improvement over time can also be measured. As the improvement rate moves towards 0, this coverage and accuracy metrics indicate the expected coverage rate and accuracy of the tool on unseen texts. The corollary of this is that upon each iteration of the development and testing cycle, the time per cycle will decrease (there will be less missing stems to add, less phonological or morphotactic rules to add or revise).

5.2 *Structures*

In this section, I briefly review some structures and statistics of the system in its current status. From a high-level perspective, I discuss the linguistic analyses which were implemented in the system to date.

5.2.1 *Lexical classes*

An overview of the treatment of Lushootseed morphophonology implemented here begins with the collections of morphotactic primitives, stored as `lexc` lexicon files and containing the roots, affixes and clitics to be combined in word formation. Some of these lexical classes are mnemonically labeled for a semantic or morphotactic coherence such as `noun-root`, `verb-root`, `constant` (i.e. no morphotactic potential), `aspect`, `verb-pc-suff-5`. While this classification serves as a lower bound for the granularity of morphotactic procedures, several morphotactic procedures target the union of one or more of the basic classes. Reduplication, for example, occurs on both noun and verb roots. In the current model, clitic hosting and the hosting of lexical suffixes can occur on a stem built from any of the three root classes. From a high level, the system can be seen as three pipelines (noun, verb, constant) which converge to define the network of `BaseWords`, the

-1	0	1	2	3	4	5
<i>dx^w(s)</i> - ‘CTD’	$\sqrt{R}_{(V)}$	<i>-il</i> ‘INCH’	<i>-b</i> ‘MD’	<i>-tx^w</i> ‘ECS’	<i>-yi-</i> ‘DAT’	<i>-t</i> ‘ICS’
			<i>ag^wwil</i> ‘AUTO’	<i>-dx^w</i> ‘DC’	<i>-bi-</i> ‘MAP’	<i>-c/-s</i> ‘ALTV’
			<i>alik^w</i> ‘ACT’		<i>-di-/i-</i> ‘SS’	<i>-b</i> ‘CSMD’
			<i>-ilut</i> ‘PRPV’			
			<i>-ət</i> ‘INCRP’			
			<i>-ab</i> ‘MTHD’			

Table 5.2: Affixal template of Lushootseed verb stems given in Beck (n.d., 30).

latter class serving as a cyclic attachment point for pro- and enclitics. Seen this way, it’s notable that the pipelines for noun and verb cross in certain places. This is in correspondence with the morphs which verbalize noun stems and others which nominalize verb stems. The next section on morphotactics gives further information on the combinatorics of these elements.

Phonological rules, on the other hand, cross-classify the morphotactic types, targeting collections of roots and/or stems based on orthographic environments. So, on the whole, the morphotactic organization into basic roots, affixes and clitics can be seen as merely a single dimension within a multidimensional system. Following the morphotactic overview, I provide an overview of the phonological rule system which is composed with the morphotactics networks in multiple stages.

5.2.2 Morphotactics

The current morphotactic system has been put together by interpretation of the information in Beck (n.d.), and by experimentation with the data. The information in Beck’s grammar includes descriptive paragraphs about particular affixes and their properties, as well as an overview table shown for verbs (reproduced here as Table 5.2). My experimentation with the data has added to and modified the template given in Table 5.2 and derived templates for nouns and constants as well.

In addition to the view of the system as three pipelines (noun, verb, and constant), the building up of a Lushootseed word can, from another perspective, be seen as adding successive layers

to a core. At the innermost level, the process begins by loading NounRoot and VerbRoot lexicons and applying optional reduplication to both. Reduplication, while not shown in Beck's template, provides both prefixal and suffixal material to the roots.

After reduplication is attached and computed, the prefixal layer is added. Looking at the verbal template of Table 5.2, position class 1 only contains $dx^w(s)$ - (CTD). In my system, this class has been modified to also contain the $?i\dot{l}$ - (PRTV), which is not included in the template, as well as the prefixes of motion tul - (CNTRFG) and $li\dot{l}$ - (PRIV). Because PRTV also attaches to nouns, this prefix class is optionally applied to both the noun and verb stems that are the result of the optional reduplication step. To the left of position class 1 (of Table 5.2), nouns can either take the 'propriative' prefix, which denominalizes them, or they can take a possessive prefix. For verb stems, an aspect prefix may occur. The aspect prefix can occur on verb stems derived from a nominal base by application of the propriative. In the XFST source code, these dependencies are carried out by a single regular expression for each of the major classes: Noun and Verb, shown in (30). Note the addition of verbalized noun-stems (NPostRed) to the verbal pipeline.

Next, the morphotactic system turns to suffixation. The noun stems are finalized by allowing possessive suffixes to attach, this is shown in (31).⁷

(30)

```
# go to lev1, add prefixes to nouns and verbs
define VLev1 (Asp) [ [ (VPCp1) VPostRed ] | [ Prop (VPCp1) [ Except .P. NPostRed ] ]
    ];
define NLev1 (PoPref) (VPCp1) [ Except .P. NPostRed ];
```

(31)

```
define NStems NLev1 (PoSuff);
```

⁷Ideally, the possessive suffix and possessive prefixes should be constrained to occur in complementary distribution since together, they complete a paradigm. However, this constraint has not yet been implemented.

The suffixation of verbs, on the other hand, has quite a bit of structure. Looking again at the template of Beck (n.d., 30) (reproduced in Table 5.2), there are a number of constraints and extensions which should be noted. First, the secondary suffixes of position class 4 are only licensed when a form from position class 5 occurs. Second, there is a class of object markers which is not indicated in the template. These markers occur outside of position class 5 and they require some sort of transitivizer. That is, the object markers require either a marker from position class 3 (the external or direct causative), or a marker from position class 5 (the internal causative, causative middle, or allative applicative). These dependencies can be summed up by two XFST regular expressions (shown in (32)), the first of which allows the verb stem to be optionally suffixed by position classes 1 or 2, the second capturing the dependencies between classes 4 and 5, 3 and 5, and the object markers.⁸

At this point the shells of reduplication, lexical suffixation, prefixation and other suffixation have been applied. These forms can be grouped together as an inflectional word (IWord) after attaching including the optional attachment of a nominalizing proclitic to the verb stems⁹ and then composing the first stage of phonological rules, shown in (33). The composition operator of XFST is `.o..`

(32)

```
define VLev2 VLev1 (LSuff) (Inch) (VPCs2);
define VStems VLev2 ([[Cause|[(Second) VPCs5]] (Obj)]) (Rel);
```

(33)

```
define Baseword NStems| (Nmlz) VStems;
```

⁸Note that in the XFST source files of the implementation, position classes 1, 3 and 4 of Table 5.2 were renamed ‘Inch’, ‘Cause’, and ‘Second’, respectively.

⁹It’s convenient to attach the nominalizing proclitic here because (1) it doesn’t need to apply more than once (it can spin against an *s*-deletion phonological rule), but the proclitics are allowed to cycle in general and (2) the attachment of possessive clitics to nominalized verb stems depends on the presence of these nominalizers.

```

## do phonology against words
source phon.foma
define IWord Baseword .o. Stath
                        .o. FinalHard1
                        .o. FinalHard2
                        .o. ?AH
                        .o. SchwaIns
                        .o. FVREM
                        .o. TSCombine
                        .o. LS2S
                        .o. SHREDUCE
                        .o. LDREDUCE
                        .o. VVC
                        .o. ?əbs;

```

Finally, the non-inflecting, or constant lexical items are combined with another special suffix, the ‘desiderative’¹⁰ and the resulting network, *Word*, is combined with the Kleene closures of the pro- and enclitic networks, shown in (34). Notice the final XFST statement, *regex*, places the resulting network in memory ready to be queried or written out in a storage format to be used later.

(34)

```

define DSD [ {-ab} ];
define Word [IWord|Const] (DSD);

## attach clitics and do clitic phonology
read lexc proclit.lexc

```

¹⁰The desiderative behaves like an enclitic with respect to its effect on phonological processes, but is glossed and discussed as a word-level formative in the literature and in the data. Therefore, in the current implementation, I attach it after the word level phonology has occurred.

```

define ProClit;
read lexc enclit.lexc
define EnClit;

regex ProClit* Word EnClit* .o. ?REM
                                .o. UREM
                                .o. SCHWADEL
                                .o. VREDUCE1
                                .o. VREDUCE2
                                .o. SDREDUCE
                                .o. HEPEN
                                .o. ConSync
                                .o. VLEN
                                .o. əs
                                .o. RemoveMarkers ;

```

In summary, the system can be viewed as intersecting pipelines of derivation and inflection applying to the base lexical classes of noun, verb and constant, or it can be seen as a series of the successive layers of reduplication, prefixation, lexical suffixation, other suffixation and cliticization applied to a core. On the latter view it's important to note that some of the layers apply equivalently to nouns and verbs (reduplication) while other layers do not. Finally, the image of Figure 5.3 shows the resulting morphotactic network for a Lushootseed word. This image shows something of a skeleton of the final network, which has granularity down to the character level. Additionally, this image does not show the results of composition with the phonological component (because those changes would have effect on a cross-classification (i.e. the orthographic environments) of the items within lexical classes, and that level of detail is not represented in this image.

Essentially, the Beck's template (Table 5.2) was adopted nearly wholesale but with the following changes and extensions: (1) reduplication procedures were implemented cyclically near-

est the root with CV- reduplication occurring outside of CVC reduplication; (2) lexical suffixes were allowed to attach just outside reduplication (but inside all other affixation); (3) inflectional affixes were added outside of the derivational material (for verbs: nominalizers and aspect prefixes, object marking and the relational suffix; for nouns: possessive affixes); (4) furthest out, proclitics and enclitics were allowed to attach (the former cyclically). The network graph in Figure 5.3 presents the current implementation’s morphotactic system in a representation in which each position class is a single node in the network. In fact, the actual network explodes each of these nodes into subnets based on the lexical material contained within them.

5.2.3 Phonological Rules

The current system implements 23 phonological rules, with 12 applying at the word level (before clitics are attached) and 11 applying “post lexically”. Among the word-level rules are voicing of final consonants and schwa-insertion procedures. The examples in (35) illustrate the word-level application of final voicing.

- (35) a. *g^wəl tu^xw ʔulila^xad k^wi ʔudsk^waʔtubu^ləd*
g^wəl tu^xw ʔu=lil•a^xad k^wi ʔu=d=s=k^waʔ-t-ubu^ləd
 SCONJ just IRR=far•side REM IRR=1SG.PO=NM=released-ICS-2PL.OBJ
 ‘And I’ll just let you guys go on the far off side.’ Seal Hunters (ML)
- b. *ʔu^xəd^yi^d k^wi ʔu^lčil ʔə tiʔəʔ adsčistx^w čx^wa*
ʔu=čəd-yi-t k^wi ʔu=ʔčil ʔə tiʔəʔ ad-sčistx^w čx^wa
 IRR=push.aside-DAT-ICS REM IRR=arrive PR PROX 2SG.PO-husband 2SG.COORD
ʔuʔə^ltx^w ʔə tiʔi^l
ʔu=ʔə^l-tx^w ʔə tiʔi^l
 IRR=be.fed-ECS PR DIST
 ‘Set it aside for when your husband arrives and you will feed him with it.’ Seal Hunters (ML)

- c. *λuhuyudəx^w* *əlg^wə? dx^w?al* *k^wi g^wəstitičs*
λu=huyu-t=ax^w *əlg^wə? dx^w-?al* *k^wi g^wə=s=ti-tič=s*
 HAB=made-ICS=NOW PL CNTRPT-at REM SBJ=NM=ATTN-get.cut=3PO

‘They always prepared it until it was cut into small pieces.’ Seal Hunters (ML)

In (35a), the internal causative, /-t/, surfaces unvoiced because there are further word-level suffixes to its right. In (35b), the [-d] allomorph surfaces, because the internal causative is in final position. In (35c), the [-d] allomorph shows up even when an enclitic follows, demonstrating the need for the distinction between word-level phonology which occurs before clitic attachment. That there is another set of rules which apply at the clitic level was shown in (29), where presence of a glottal stop on the realization of aspect prefix is dependent on enclitic attachment.

It should be noted that these rules vary widely in their domain of applicability. For instance, there is a rule which removes the initial *h* of *hay* (to make) when preceded by the stative marker (36a), which does not affect other words with similar phonological shape (36b).

- (36) a. *x^wi? g^wəsasaydubs* *?ə ti?ə? di?ə?*
x^wi? g^wə=s=?as-hay-dx^w-b=s *?ə ti?ə? di?ə?*
 NEG SBJV=NMLZ=STAT-known-DC-PASS-3.POSS PR DET here

‘It is not known by the children.’ Basket Ogress (MS)

- b. *?əshuyəx^w* *ti?i?ə?*
?as-huyu=ax^w *ti?-i?ə?*
 STAT-made=NOW DSTR-DEM

‘They are ready.’ Basket Ogress (MS)

Other rules apply very broadly, such as a rule which allows any vowel to be lengthened for dramatic emphasis. Still other rules were implemented but allowed to apply optionally. For example, the verb *pusu* sometimes appears with, sometimes without its final vowel, even in a

similar environment, as illustrated in (37).¹¹

- (37) a. *x^wi? g^wəspusdubs* *ti?ə? di?ə? kikəwič*
x^wi? g^wə=s=pusu-dx^w-b=s *ti?ə? di?ə? ki-kəwič*
 NEG SBJV=thrown-DC-PASS-3.POSS DET here ATTN-hunchback
 ‘Little Hunchback isn’t hit.’ Basket Ogress (DM)
- b. *ləpuspusutəb*
lə=pus-pusu-t-b
 PROG=DSTR-thrown-ICS-PASS
 ‘He was being thrown at.’ Basket Ogress (DM)

Other rules are theoretically uninteresting but important for the practical nature of an implementation of the system. One such rule cleans up boundary symbols after all other rules have applied. Most rules fall between these extremes, applying generally but in specific phonological environments.

In classifying and organizing the phonological rules according to order and domain of application, it should also be kept in mind that the final object will be a transducer which is composed with and which will cross-classify much of the structure in an extant network (the morphotactics), rearranging nodes and connections according to network composition algorithms as well as optimization procedures implemented in the FOMA libraries, down to a character level granularity. This is to say that as with the discussion of the morphotactics system above, the description in the source files is only one way to examine the final object. To see this, consider for one, that the descriptions in XFST and lexicon source files are ambiguous with respect to the final network. There is more than one way to define an equivalent transducer. Also, the number of rules in a network is partly determined by what counts as a rule. To exemplify with the current model, there is a set of rules which delete equivalent consonants across a morpheme boundary. In the

¹¹Beck (p.c.) points out that treating this rule as optional misses an important generalization. In fact, a broader survey of the data reveals that the final vowel of CVCV roots is consistently deleted before the DC suffix.

source, individual rules for each consonant to which the rule can apply are written. They are then composed together as a consonant deletion rule which is later composed into the morphotactic network at the post-word level. So, in looking at the immediate part of the code in which the post-lexical phonological system is composed, there is only the single rule representing this entire family of deletions. In order to rigorously count rules, a base granularity would have to be established. In general this discussion is intended to emphasize that particular organizations within the source can be seen as engineering decisions whose construction criteria were drawn as much from needs of organization and maintenance as linguistic analysis. Because of this, and the intractability of examining the resulting network by hand, the system's input and output have to be the primary evaluation measure for the system.

Nevertheless, the phonological transducer can, in principle, be visualized, as was done for the morphotactic network in Figure 5.3, using the classes of inputs as individual arcs. However, upon construction of the transducers for word-level and post-word-level phonology I found that the former has 390 states with 7103 arcs and latter has 240 states with 4589 arcs — far too many for me to usefully render in a visualization.

5.3 Relationship to Descriptive/Documentary Efforts

Thus far, the analysis of the first three texts in Beck and Hess (2014, 2015) has led to the repair of a number of typographical errors,¹² some discovery of ambiguity patterns henceforth unnoticed as well as a subtle rule interaction which falls out of the forms used on the orthographic and morphophonemic lines of Beck and Hess' IGT.

To briefly illustrate ambiguity discovery, the system (correctly) produces two readings for the form $\text{ʔuʔu}\check{x}^w$. In one reading a perfective prefix $/\text{ʔu-}/$ is affixed to the stem $\text{ʔu}\check{x}^w$ ("go"), in the other diminutive reduplication is prefixed. In personal communication, fluent speaker Zalmal Zahir commented that although he could not recall ever hearing the diminutive reduplication attached to the verb $\text{ʔu}\check{x}^w$, he considered it to be a legitimate wordform, and added that the two

¹²This is reported in order to highlight the utility of machine-implemented analysis, not to belittle the careful work of the researchers who created the data.

forms would be differentiated through stress.¹³ This example serves to illustrate some of the ways that an implemented analysis can be of aid to a linguist — it can highlight consequences of the analysis which are easily overlooked because they are less salient to the linguist who is familiar with the language.

Another, similar example of such an interaction was found in analyzing forms in which a perfective prefix is completely deleted in the context of a temporal proclitic, as in (38).

- (38) *g^wəl tu∅səg^wqtag^wəl tiʔə? cədiʔ tasčəbaʔtəb,*
g^wəl tu=ʔu-səg^wq-t-ag^wəl tiʔə? cədiʔ tu=ʔas-čəbaʔ-t-b
 SCONJ PST=PFV-whisper-ICS-RCP DET s/he PST=STAT-backpack-ICS-PASS
ʔəsqiqtub
ʔas-qiqt-tx^w-b
 STAT-confined-ECS-PASS

‘Those that had been carried and confined whispered to each other.

Basket Ogress (DM)

Note that this deletion fits with other observed alternations. It is already captured by the rules in (5) and (6). The two rules conspire to first remove the glottal stop and then one of the *us*, so there is no remnant of the aspect prefix left in the surface form. The prediction, then, is that anywhere that a temporal proclitic attaches to a verb without an intervening stative prefix (the stative prefix is in complementary distribution with the perfective), two analyses are possible: one with and one without the perfective prefix.¹⁴ This anecdote, then, again illustrates the ways in which machine implementation provides the linguist with insights into both the consequences of their analyses, and the phenomena which can sometimes be “hidden in plain sight” in their own data.

¹³However, as stress is not represented in the standard orthography, the system’s analysis is taken to be correct.

¹⁴David Beck (pc, 2013) suggests that this ambiguity is spurious. Although there are many examples in the corpus of this deletion, he does not find a semantic motivation to posit the perfective in these examples (and there is no surface phonological evidence on which to rely).

5.4 Conclusion

This chapter has presented the initial morphophonological analyzer associated with the overall architecture of a text-processing system which I discussed in Chapter 4. I have presented the methodology and resources used, some initial results, and some of the linguistic organization (the classes and rules implemented). In the remainder of this document, I will turn to the development of the syntactic grammar. The next chapter describes some of the initial implementation work done for word order and morphological analysis on the syntactic side. This leads into the more detailed case-studies which appear subsequently.

Chapter 6

GRAMMAR ENGINEERING EXAMPLES

Previous chapters discussed the motivations, the theory, the technology and tools which are used in the project as well as the initial implementation of the morphophonological analyzer. In this chapter, I turn to the syntactico-semantic system, which is the topic of the remainder of this document. This chapter is included in order to provide a “warm-up” to the larger case studies of Chapter 7 and Chapter 8 by presenting two grammar engineering examples pulled from the initial development work which was done in order to get the syntactic meta-grammar off the ground and running. These examples serve to illustrate how the theories, tools and techniques described in the introductory materials above come together in practice, as well as to begin to familiarize the reader with the system under discussion in the later case-studies. The first example shows how I modified the word order library of the Grammar Matrix customization system to support the word order phenomena of Lushootseed. The second example presents the strategy undertaken for capturing the distribution and semantics of Lushootseed object marking morphology.

6.1 VSOv2 Word Order

After the gathering of linguistic resources and the preparation of the data into testsuites (see §4.7), the next step was to implement a first analysis. Because the Grammar Matrix customization system I had forked is a grammar generation system (see §4.3), I began by filling out the questionnaire to see just how far the system could take me “out-of-the-box”.

One of the initial choices one makes when using the customization system regards major constituent order, referred to in the system as word order. Based on the input choices file, the customization system provides an output grammar with phrase structure rules which will constrain the potential orderings of subject, verb and object. The current grammar generation

code which provides a generalized word order library for the customization system was carried out by Antske Fokkens (Fokkens, 2010) as an update to the system built by Emily M Bender and Dan Flickinger (Bender and Flickinger, 2005). Fokkens' system (at the time of my fork) allows user-linguists to choose between ten options: the six logically possible arrangements of S, O and V; V-initial or V-final; free (pragmatically determined); or V2 (finite or auxiliary verb in second position, with free word order otherwise). The grammar generation code, i.e. the word order library, will include a particular set of types implementing head-valence rules according to the user's choices.

For someone with a background only in “pen-and-paper” HPSG, it may be surprising to find that the grammars output by the Grammar Matrix customization system include binary and unary rules only, because, phrase structure rules with ternary (or greater) arity are common within theoretical HPSG. The reason for this binary-only rules is a technical one. Higher arity rules can always be factored down into a series of binary rules with no loss of information (Chomsky, 1963, 371) and the result of this factoring down is a lower complexity search space for parsing.¹

To construct a simplistic example: consider an SVO language without auxiliaries. In the traditional, VP-centered analysis, two rules will be required: a head-complement rule for combining a verb and its object along with a subject-head rule for combining the VP with its subject. For a verb with one argument to the left and another to the right, factoring the dependencies into two rules is probably the intuitive approach for most people. However, when building a di-transitive VP from two complements, many theoretical HPSG analyses would combine the verb with both complements according to a single, generalized, phrase-structure schema. An alternative approach is to use the recursion application of a single head-complement rule to combine each argument with the verb, one at a time. The latter is the approach implemented in Matrix-derived grammars. One trade-off of using binary-only rules is that there are more situations where, unless care is taken, a pair of rules may apply in either order, leading to spurious ambiguity.

¹Jurafsky and Martin 2009, Ch. 13 provides an introduction to parsing as a search problem.

The other aspect of major constituent ordering which the word order library treats regards auxiliary verbs. Users can specify whether or not their language includes auxiliary verbs as well as the ordering properties of those verbs with respect to the rest of the clause.

Lushootseed word order is not directly represented in the list of supported options, however, so I extended my fork of the customization system in order to generate the rules necessary to model basic Lushootseed sentence structure. The rest of this section describes that work in a way which will become a familiar pattern for the remainder of this document: a synopsis of the phenomena to be modeled precedes a discussion of potential analyses, followed by a presentation of the details which arose in implementation and testing.

6.1.1 Lushootseed word order

From a very high level, the basics of Lushootseed word order to be implemented can be described as the interaction of three constraints which are discussed in more detail below.

- Lushootseed is VSO overall, but there is a set of grammatical items which always appear in sentence-second (Wackernagel's) position: main clause subject-markers and question markers.
- Lushootseed disallows the direct combination of a bivalent verb with two third-person noun phrases. In these 3rd-on-3rd situations, one NP or the other NP must be marked as an oblique (by embedding within an prepositional phrase). Another way to say this is to say that syntactically, Lushootseed predicates have a single argument slot which can accept a direct argument.
- This argument slot is subject to the rule of “one nominal interpretation” (Gerdt's, 1988, p. 57) written out in (39).

One of the ramifications of the first point above is that some aspects of Lushootseed word order can be modeled by the VSO option of the customization system, but this approach entails

a difficult path to capturing the second-position facts of the sentence-level clitics. On the other hand, the out-of-the-box V2 option of the customization system is intended to model a system in which the finite verb appears sentence second, but word order is otherwise free. So while that “out of the box” option provides some of the machinery for an element into second position, it isn’t a wholly appropriate out-of-the-box solution for Lushootseed.

As a typological aside, I note that the prohibition on two full noun phrases in a clause which Lushootseed enforces has reflexes in other Salishan languages as well — although less categorically: Hukari (1976, p.98) reports no such V-NP-NP constructions in a twenty-five minute oral text from Halkomelem [hur], but says such constructions do occur “spontaneously” if infrequently; for St’át’imcets [lil] (“Lillooet”) Burton et al. (2001) observe that 60 pages of text from van Eijk and Williams (1981) contain just six examples.² The Lushootseed prohibition on two full noun is probably connected to a broader Salishan phenomenon wherein 3rd-on-3rd constructions are generally dispreferred³ This typological pattern means that a Lushootseed-focused implementation may be able to form some of the basis for future generalization into a Salishan meta-grammatical system.

The third point above is best illustrated by example. The sentences in (40) serve to illustrate the effect of the rule of interpretation. Mainly, the rule provides a means of disambiguation when a transitive verb appears with a single argument. In (40a), “the child” is interpreted as the subject of *ʔuχʷ* and there is no ambiguity because the verb is intransitive and only bears a subject. However, in (40b), “the child” has to be the syntactic object. This is because the applicative affix creates an additional valence slot for an object (which indicates a role-player in the event structure towards which the subject’s motion is aimed) and once there are two potential argument slots for an NP to fill, the rule of interpretation resolves the question in favor of the syntactic object. Finally, in (40c), the child is interpreted as the syntactic subject because a 1st person singular object marker has saturated the object position. Subject is the only valence slot

²See the introduction to Gertds and Hukari (2008) for further references.

³Beck (pc) points out Gerdts 1988; Jelinek and Demers 1983 for Halkomelem and Coast Salish reflections of this phenomenon (respectively).

available at the point where *ti čáčas* is combined with *ʔuʔuχ^wcbš*.

(39) When a Lushootseed NP appears as the direct argument of a predicate, its interpretation as a syntactic object is preferred.

(40) a. *ʔuʔuχ^w ti čáčas*
 ʔu-ʔuχ^w ti čáčas
 PFV-GO DET child

‘The child went.’

b. *ʔuʔuχ^wc ti čáčas*
 ʔu-ʔuχ^w-c ti čáčas
 PFV-GO-ALTV DET child

‘Someone/something went after the boy.’

(Hess, 1995)

c. *ʔuʔuχ^wcbš ti čáčas*
 ʔu-ʔuχ^w-c-bš ti čáčas
 PFV-GO-ALTV-1sg.obj DET child

‘The child went after me.’

6.1.2 Data

One of the experimental questions at play in this work is just how closely one can hold to the existing Grammar Matrix customization system options and paradigms when implementing a grammar of a Salishan language. Therefore, I started by creating a grammar using the VSO option from the existing system. I listed the subject markers as a type of noun: they were marked as incompatible with determiners (syntactically, they started off in the lexicon fully saturated). This solution provided analyses for a few simple, intransitive sentences. However, the solution of treating Lushootseed subject markers as noun phrases fell short of the goal in two ways: it

	Singular	Plural
1st	čəd	čəł
2nd	čəx ^w	čəlap/čələp

Table 6.1: The forms of Lushootseed main clause subject markers

failed to analyze sentences in which the subject markers appear before the main predicate and it overgenerated by allowing subject markers to appear anywhere that a full NP can appear. In order to encapsulate these facts for testing, I created a testsuite of positive and negative examples related to general word order.

The forms of the Lushootseed subject markers are given in Table 6.1.⁴ This testsuite is

⁴Lushootseed does provide an optional marker to indicate the plurality of a dropped 3rd person argument. The 3rd person plural marker is not included in the table because it is neither morphologically nor syntactically connected to the paradigm under discussion. The form appears in written texts variously as *əlg^wə?*, *həlg^wə?* or *(h)ilg^wə?* and can indicate the plurality of a subject or object. The sentences below provides some illustration of its usage.

- (i) *lətələwil həlg^wə?*
 lə=tələwil həlg^wə?
 PROG=RUN PL

‘They are running.’

- (ii) *?ə(s)šudx^w čəd həlg^wə?*
 ?as-šut-dx^w čəd həlg^wə?
 STAT-SEC-DC 1SG.SUBJ PL

‘I see them.’

- (iii) *?ə(s)šudx^w čəd ti bads (h)əlg^wə?*
 ?as-šut-dx^w čəd ti bad-s həlg^wə?
 PROG-SEC-DC 1SG.SUBJ DET father-3SG.POSS PL

‘I see their father.’

Hess and Hilbert (1995b, 42.1 pg.87)

Number	Well-formed	Item	Number	Well-formed	Item
0	T	ʔuχ ^w čəd	17	T	ʔuχ ^w -c čəx ^w ti čáčas
1	T	ʔuχ ^w čəx ^w	18	T	ʔuχ ^w -c čəł ti čáčas
2	T	ʔuχ ^w čəł	19	T	ʔuχ ^w -c čəlap ti čáčas
3	T	ʔuχ ^w čəlap	20	F	čəd ʔuχ ^w -c ti čáčas
4	F	čəd ʔuχ ^w	21	F	čəx ^w ʔuχ ^w -c ti čáčas
5	F	čəx ^w ʔuχ ^w	22	F	čəł ʔuχ ^w -c ti čáčas
6	F	čəł ʔuχ ^w	23	F	čəlap ʔuχ ^w -c ti čáčas
7	F	čəlap ʔuχ ^w	24	F	χ ^w ul' ʔuχ ^w -c čəd ti čáčas
8	F	χ ^w ul' ʔuχ ^w čəd	25	F	χ ^w ul' ʔuχ ^w -c čəx ^w ti čáčas
9	F	χ ^w ul' ʔuχ ^w čəx ^w	26	F	χ ^w ul' ʔuχ ^w -c čəł ti čáčas
10	F	χ ^w ul' ʔuχ ^w čəł	27	F	χ ^w ul' ʔuχ ^w -c čəlap ti čáčas
11	F	χ ^w ul' ʔuχ ^w čəlap	28	T	χ ^w ul' čəd ʔuχ ^w -c ti čáčas
12	T	χ ^w ul' čəd ʔuχ ^w	29	T	χ ^w ul' čəx ^w ʔuχ ^w -c ti čáčas
13	T	χ ^w ul' čəx ^w ʔuχ ^w	30	T	χ ^w ul' čəł ʔuχ ^w -c ti čáčas
14	T	χ ^w ul' čəł ʔuχ ^w	31	T	χ ^w ul' čəlap ʔuχ ^w -c ti čáčas
15	T	χ ^w ul' čəlap ʔuχ ^w	32	F	ʔuχ ^w -c ti čáčas tsi čáčas
16	T	ʔuχ ^w -c čəd ti čáčas	33	T	daɣ čəd čick ^w ʔas-laɣ-il

Table 6.2: Testsuite items for subject markers and general word order constraints. The verb *ʔuχ^w* (“go”), the four subject clitics of Table 6.1, a noun phrase (*ti čáčas*, “the child”), the presence of an applicative (deriving a transitive) and one or two adverbs are permuted to check sentence second position of the subject markers. Item (32) illustrates the prohibition on two explicit NP arguments. Item (33) tests for the proper interaction between the subject clitic and multiple sentence-initial elements.

presented in Table 6.2. The goal of the testsuite is to capture the placement of subject markers in the context of transitive and intransitive verbs with and without a sentence-initial adverb. Note that the usage of the adverb *ǰʷul* (“just”) may actually present some semantic anomaly when combined with these sentences which would make them unacceptable to a native speaker. This fact is recorded in the fuller form of the [incr tsdb()] testsuite, where it is noted that these are constructed examples created by me. Nevertheless, these potentially questionable examples serve to test the general pattern which is that predicate adverbs may be sentence-initial and their presence interacts with correct placement of the subject-marker. What’s more, these examples can, over time, be replaced with attested examples from the corpus as such becomes available to me.

Based on a suggestion from Bender (pc), I began to explore an alternative approach to the placement of subject markers based on treating them not as nominals but as auxiliary verbs.⁵ Considering that the ultimate placement of subject markers is a function of sentence-level considerations, treating them as sentence-level auxiliaries, i.e. sentence-level elements, is not unnatural. Under this approach, subject markers will enforce agreement constraints for person and number on the main predicate and will saturate the valence requirement. In order to facilitate this approach within the customization system, I began exploring the potential of creating a new word order option within the word order library. When this new option is selected, the overall word order will be VSO, but the output grammar will include rules for auxiliary verbs which must appear in sentence-second position — the Lushootseed subject markers can then be indicated as such verbs in an input choices file.

⁵Beck (p.c.) pointed out that calling these elements auxiliary verbs is descriptively awkward because they don’t directly behave like lexical verbs (they can’t take morphology associated with lexical verbs, they don’t have the same syntactic distribution). On the other hand, they do play an auxiliary role for verbs (and predicative nouns, see Chapter 8), so perhaps “verbal auxiliaries” is a better term of art. I think this suggestion is worthwhile. I will continue to refer to them as auxiliary verbs in this document, however, mainly because in the implementation I’m describing here, these elements share a head-type with lexical verbs and structures in the implemented grammar which are satisfied by the head-type requirement *verb* are satisfied by these elements.

6.1.3 *Analysis and Implementation*

In terms of writing code for the new VSOv2 word order option, my approach was to leverage as much of the existing VSO implementation as possible. This provides the initial scaffolding upon which the constraints on auxiliary verb placement must be added. I also took certain aspects of the existing auxiliary verb treatment from the free word order choice. This section presents the structures I adopted and those which I adapted. It also presents the trade-offs I considered when weighing two alternative approaches to building the constraints that ensure the correct sentence-second position of the auxiliary verb.

I begin with the algorithm of the existing word order library, which proceeds by first setting the head-dependent order of head-subject and head-complement rules for a given configuration. Under a binary analysis, VSO is consistently head-first. The core-grammar contains type-definitions which generically handle linearization on headed rules, so the customization routine only has to include the relevant supertypes on a head-subject-rule, or a head-complement rule. In the VSO case, both head-subject and head-complement rules will be set to inherit from the supertype `head-initial` (41).⁶

```
(41)  head-subj-phrase := decl-head-subj-phrase & head-initial.
      head-comp-phrase := basic-head-1st-comp-phrase & head-initial.
```

To enforce that subjects attach below (before) complements, the rule type for a head-comp phrase is further constrained to have an empty `SUBJ` list (42).

```
(42)  head-comp-phrase := decl-head-comp-phrase & head-initial &
      [ HEAD-DTR.SYNSEM.LOCAL.CAT.VAL.SUBJ < > ].
```

⁶Because the grammar engineering framework at use here has been in development for many years, the existing typology provided by the Grammar Matrix core grammar provides a lot of detailed specifications out-of-the-box. Because of space constraints, I won't always be able to explain in full every supertype which appears in examples. I have opted not to redact such unexplained supertypes from the examples in this document because the details thereof are available to the interested reader via the open-source distribution of the Grammar Matrix project.

These are the main bits of the existing VSO implementation which forms a starting point for creating a VSOv2 analysis for Lushootseed. The next step is to add a pair of rule-types for combining a verb with an auxiliary. When a grammar contains auxiliaries which attach to their verbal heads in a linear order which is non-harmonic with the order of main verbs to their complements — such as, for example, the free-word order option or the SVO option when auxiliaries attach to the right of the main verb — the existing library creates attachment rules such as those in (43). These rules inherit from `basic-marker-comp-phrase`, a type defined in the Matrix core grammar which is like a head-comp phrase, but which doesn't enforce token identity of the mother's head feature with that of the head daughter, allowing the resulting head value to incorporate features from either daughter's head. The types `marker-initial-phrase` and `marker-final-phrase`, analogously to the `head-initial` and `head-final` supertypes of (41), serve to constrain the linear order of their arguments.

- (43) a. `comp-aux-phrase := basic-marker-comp-phrase & marker-final-phrase & [SYNSEM.LOCAL.CAT.VC #vc, MARKER-DTR.SYNSEM.LOCAL.CAT.HEAD verb & [AUX +], NON-MARKER-DTR.SYNSEM.LOCAL.CAT [HEAD verb, VC #vc]]`.
- b. `aux-comp-phrase := basic-marker-comp-phrase & marker-initial-phrase & [SYNSEM.LOCAL.CAT.VC #vc, MARKER-DTR.SYNSEM.LOCAL.CAT.HEAD verb & [AUX +], NON-MARKER-DTR.SYNSEM.LOCAL.CAT [HEAD verb, VC #vc]]`.

The `AUX` feature which appears in (43) is defined on verbal heads and indicates, unsurprisingly, a verb's categorization as lexical or auxiliary. It is lexically stipulated and often used, as in these examples, to constrain the application of rules to one class or the other. `vc` is mnemonic for “verbal cluster”. This feature stems from the argument-composition approach (Hinrichs and Nakazawa, 1990) to particular word orders involving auxiliaries which is used in the word-order library. None of the details about `vc` are important for the analysis presented here but my rules do inherit these constraints, because of my intention to remain as compatible as possible with

the existing Matrix-based treatment of other word-orders.

Starting from these rules, the next requirement is to ensure the proper distribution of the two auxiliary rules of (43) in such a way that the auxiliary always ends up in sentence-second position. I considered two approaches, one based on the Edge Feature Principle of Miller (1992), the other based on the *xmod* hierarchy which tracks the existence and placement of modifiers using a technique pulled from the ERG (Flickinger, 2000).

EFP for subject-auxiliary placement

Edge-related phenomena are not rare across languages. Even languages such as English, which rely on word order for grammatical function marking present some edge sensitivity in constructions such as the Anglo-Saxon genitive. The idea of an Edge Feature Principle (EFP) enforcing the propagation up of edge-related dependencies and their satisfaction was introduced in HPSG at least as early as Miller (1992). Tseng (2003) broke Miller's *EDGE* features out into left and right variants for keeping track of liaison phenomena in French. Crysmann (2010) demonstrates the utility of the EFP outside Indo-European in an analysis of Hausa [hau].

Sanghoun Song, in connection with his work on information structure, added two features to the core-grammar which are directly related to edge phenomena (Song, 2017). Song added the luk-valued⁷ features *L-PERIPH* and *R-PERIPH* to the type *synsem* and uses them in his library to enforce constraints on clause-initial and clause-final focus. With this in place, all that is needed to create an implementation of Miller's EFP is to ensure the proper propagation of these features in binary rules. This can be done by adding the relevant constraints to the type *basic-binary-phrase* as shown in (44).

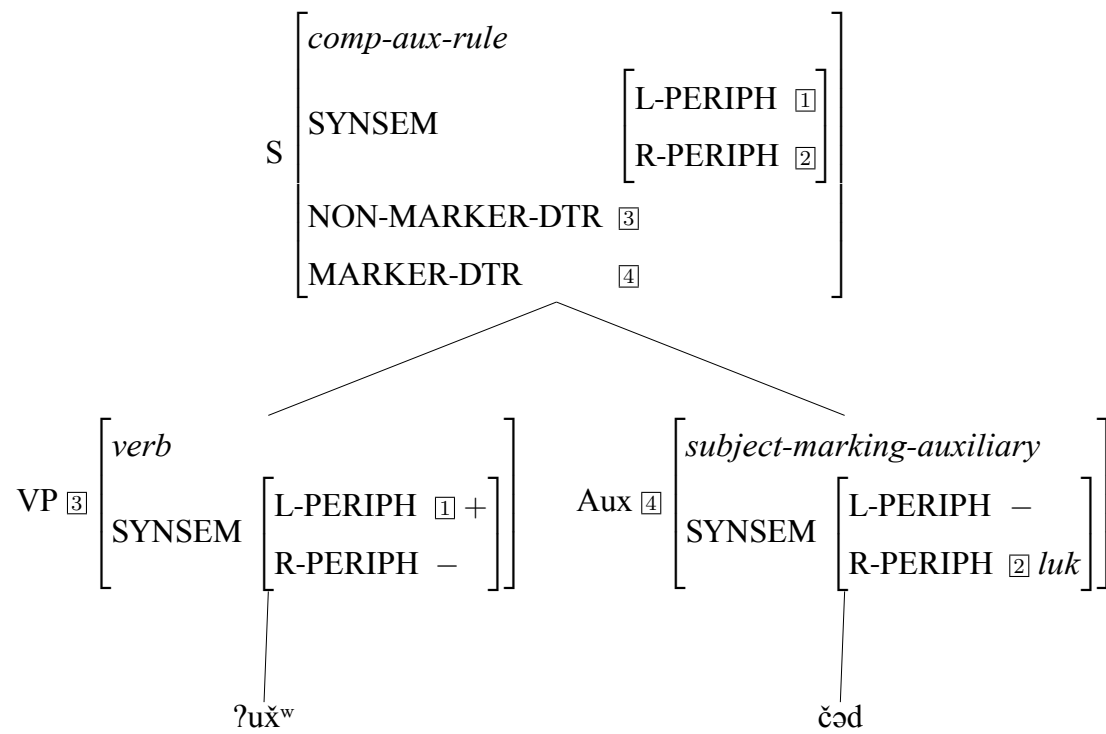
⁷The three-valued logic of Jan Łukasiewicz (1970) inspired this eponymous type-name which is commonly used in the Matrix core grammar in many situations which could be potentially handled by a simple boolean. The example under discussion, that of the *PERIPH* features, illustrates the utility of the third value: *synsems* start out in the lexicon with the value *na* on their *PERIPH* features because the lexicon cannot presuppose, for most elements, whether they will or will not appear on the periphery. Then, grammatical elements which do care that an element be on the periphery (for example) can specify a supertype which subsumes *na* and +.

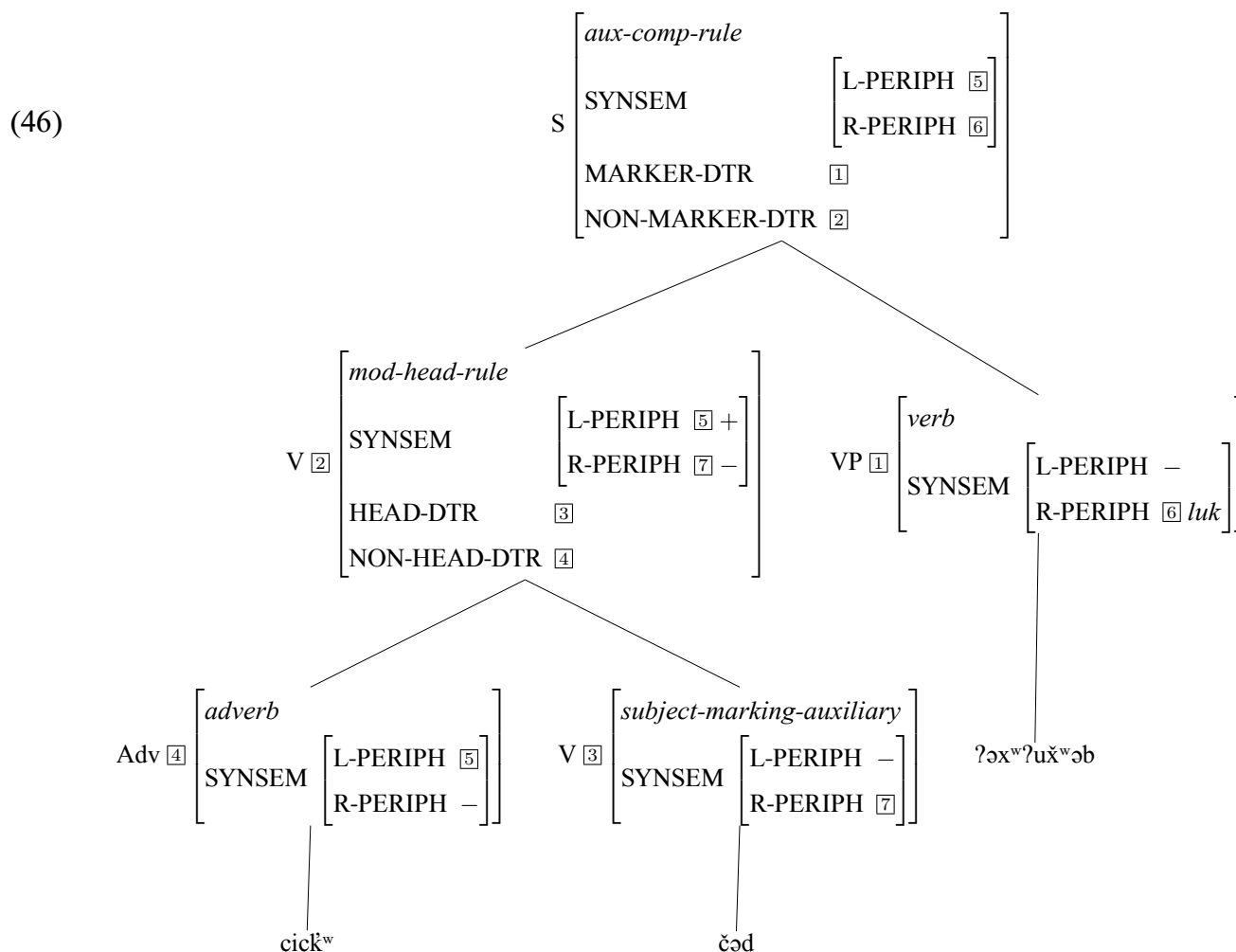
(44) `basic-binary-phrase` :+
 [SYNSEM.L-PERIPH #lp,
 SYNSEM.R-PERIPH #rp,
 ARGS < [SYNSEM.L-PERIPH #lp,
 SYNSEM.R-PERIPH -],
 [SYNSEM.L-PERIPH - ,
 SYNSEM.R-PERIPH #rp] >].

An EFP based approach would handle the construction where the subject auxiliary attaches to the right of the main predicate (schematically: *V, aux, O*-type sentences) by constraining the L-PERIPH feature of the predicate (the complement of the auxiliary) to +. This constraint can be expressed on the *comp-aux-rule* (43a) which combines these elements. An example phrase structure tree showing how the PERIPH constraints of (44) fall down onto the constituents of a simple, binary phrase built with (43a) is shown in (45). This example includes the additional constraint just mentioned, that the NON-MARKER-DTR, the main predicate, be constrained to L-PERIPH +.

The complementary construction in which the subject-marking auxiliary should be picked up on the left of the verb, will be realized in two rules. The first combines an adverb (or some left-adjoining element) with a subject marker on its right. This rule must be written to constrain the adverb's L-PERIPH feature to +, so that only an element on the left periphery may combine with a subject marker. The resulting *mod-aux-phrase* can then go through a rule which will allow it to combine with a predicate on its right. The phrase-structure tree in (46) shows how these constraints are realized in a minimal construction with two rules (both of which inherit from the EFP constraints of (44)): a *mod-aux-rule* constrained as described above, and a *aux-comp-rule* as shown in (43).

(45)





This succinctly illustrates the positive part of the EFP-based analysis. However, as described so far, this analysis is incomplete. It provides a mechanism to license the positive examples from Table 6.2, but there is no constraint to rule out the negative items of (4)–(11). Items (4)–(7) have a single subject-marking auxiliary on the left of the main predicate. Without further modification, the lexical subject marking auxiliaries will be able to pass through the *aux-comp-rule*. Recall that the EFP, as implemented in (44), would simply mark the left daughter of a binary rule as R-PERIPH – and the right as L-PERIPH –. The putative *aux-comp-rule* which is intended to combine a *mod-aux-phrase* with a predicate can also apply to a lexical auxiliary which has not combined with a left-peripheral element. Similarly, in (8)–(11), the modified verb is compatible with the proposed L-PERIPH + constraint on the *comp-aux-rule*. The crucial observation here is that the

attachment of auxiliaries to the left of the predicate should only be allowed when they have already gone through some phrase structure rule and the converse is true for lexical auxiliaries: they should only attach to the right (of the main predicate). In order to rule out the negative examples, the analysis must be extended to capture a lexical/phrasal distinction.

Given the influence of the Lexical Integrity Principle withing HPSG (Bresnan and Mchombo, 1995) and DELPH-IN, it is unsurprising that the Matrix core-grammar provides a lexical/phrasal distinction within the *synsem* type hierarchy (47).⁸ The particular types of reference are *phr-synsem* and *lex-synsem*. These types are also distinguished by a feature value: *phr-synsems* are LIGHT – while *lex-synsems* are LIGHT +. The goal of the feature LIGHT is to provide a mechanism which allows particular heads to retain some properties of a lexical-like element even after combining with a complement (Abeillé and Godard, 2001).⁹

⁸Some of the other distinctions available in the full *synsem* type-hierarchy have been omitted from this listing.

⁹Using LIGHT as a stand-in for phrasal in languages which do not exhibit light phrases keeps the grammar comparable with those that do.

```

(47) synsem-min := avm &
      [ OPT bool,
        OPT-CS cog-st,
        LOCAL mod-local,
        NON-LOCAL non-local-min ].

synsem := synsem-min &
      [ L-PERIPH luk,
        R-PERIPH luk ].

; The feature LIGHT is used to model phenomena which distinguish
; shorter words from longer ones, or constituents which are lexical
; or nearly so from larger phrases. It is inspired by the work
; of Abeille and Godard (e.g., 2003).
lex-or-phrase-synsem := synsem-min &
  [ LIGHT luk ].

expressed-synsem := synsem.
canonical-synsem := expressed-synsem &
  [ MODIFIED xmod ].

lex-synsem := canonical-synsem & lex-or-phrase-synsem &
  [ LOCAL local-min,
    LIGHT +,
    LKEYS lexkeys ].

phr-synsem-min := canonical-synsem & lex-or-phrase-synsem &
  [ LOCAL local-min ].

phr-synsem := phr-synsem-min &
  [ LIGHT - ].

```

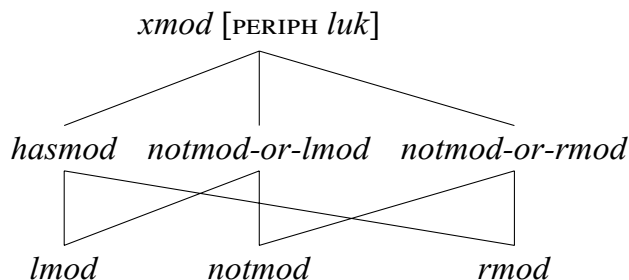
The *synsem* hierarchy, then, provides the features and types needed to flesh-out the EFP-based analysis. The *aux-comp-rule* which attaches a subject-marking auxiliary to the left of the verb must have a `MARKER-DTR` (the auxiliary) which is phrasal, or `LIGHT -`. Conversely, the *comp-aux-rule* which attaches the subject marking auxiliary to the right will require that the predicate is compatible with *lex-synsem*, by constraining its `LIGHT` value to `+`. Using the constraints on the feature `LIGHT` is preferable, for now, to constraining the *synsem* type directly because it provides for the future case when some phrasal item may be small enough in “weight”

to count as one element for the purposes of determining sentence-second position.¹⁰ The next subsection considers an alternative to the EFP-based analysis described above.

MODIFIED for subject-auxiliary placement

The second approach to word order constraints which I considered is based on the *xmod* hierarchy, which is presented graphically in (48). This type hierarchy, found in the Matrix core-grammar, is inherited from the English Resource Grammar (Flickinger, 2000). It was used originally to constrain the attachment order of pre- and post-head adnominal modifiers (Emily Bender, p.c.). Within the customization system’s Word Order library, Antske Fokkens adopted the *xmod* hierarchy for a similar purpose: the prevention of spurious ambiguity by constraint of the attachment order — but now within the domain of major-constituent orders as opposed to only modifier attachment on nominals.¹¹

(48)



Building a sentence-second positional constraint from this hierarchy is relatively straightforward given the basic word order scaffolding provided above. Since the phrase-structure rules

¹⁰Although this feature was originally proposed for use in analyzing French adverbs (Abeillé and Godard, 2001), it has proven useful in the context of unrelated languages. See Crowgey and Bender 2011 for an example from Basque, Poornima and Koenig (2009) for Hindi.

¹¹Fokkens left the following comment in the Grammar Matrix customization system source code:

This solution adopts the *xmod* hierarchy to enforce right-first attachment. That is, all arguments appearing to the right of the verb must attach before all arguments appearing to the left. The linguistic prediction of this analysis is that free word order languages do not have a consistent VP constituent, even when the verb and object are adjacent (OV order).

generate strictly predicate-first ordering when there are no auxiliaries or modifiers, the subject-marking auxiliary should attach to the right of the lexical predicate. Conversely, since the only thing which may precede the predicate are sentential adverbs, the subject-marking auxiliaries should attach before the predicate in these situations. In fact, as shown by (33) from the test-suite, the subject-marking auxiliary should attach to the right of the first such modifier.

As long as the value of `MODIFIED` is constrained to *notmod* on lexical items and properly set to *hasmod* upon the application of any adverb then this feature provides the information necessary to implement the requirements listed above. Like the EFP-based analysis, this one will place constraints on construction-specific rules. The first construction, in which the verb is initial position, adds constraints to the putative *comp-aux-rule* of (43) such that the `NON-MARKER-DTR` (the complement) must be `MODIFIED notmod` — doing this both rules in the positive examples of (0)–(3) and (16)–(19) and also rules out the negative examples of (8)–(11) and (24)–(27).

The contrasting construction, with the subject auxiliary to the left, as in the EFP-based analysis, will, minimally, involve two rules. The *aux-comp-rule* will require that the `MARKER-DTR` (the auxiliary) constituent is `MODIFIED hasmod`. As in the EFP based analysis, the rule which attaches a modifier to a subject marking auxiliary will require a constraint that the modifier is lexical: `MODIFIED notmod`. This latter constraint ensures that the subject-marking auxiliary attaches after the first modifier (in sentence-second position, as in (33) from the test-suite). This constraint on the *aux-comp-rule* also rules out the negative examples of (4)–(7).

Discussion

The next question is which of these systems is to be preferred for the task of constraining the position of subject markers, or whether the tradeoffs between them leave the matter unclear enough to maintain both analyses in the metagrammar for the time being. *Prima facie* each of the alternatives bears promise: the *xmod* hierarchy has been adapted for usage outside of strict modification scenarios (Fokkens used it while laying out analyses for major constituent orders) so using this hierarchy to constrain major constituent orders in Lushootseed is congruent with previous work in the context of Matrix-derived grammar engineering projects. The `L-PERIPH`

feature, on the other hand, has supporting theoretical literature behind it. Descriptively, there is an element of transparency — the name of the feature directly bears reference to the relevant point of reference: the left-edge of the sentence.

The two options are similar in some ways: both use two construction-specific rules to partition the data into those in which the subject-marker leads the predicate and those in which it follows and a third rule to place contextual constraints on modification of the auxiliary. Both require the notion of an opposition between lexical and phrasal constituents in order to rule out ill-formed examples from the testsuite.

They differ in that the MODIFIED-based analysis does all of its work with a single feature and its value hierarchy. Instead of drawing on notions of what's on the periphery in addition to the notion of what's lexical as opposed to phrasal, it relies on the phrase structure rules which manage the overall order to fix what occurs on the periphery and hinges on the fact that items which are (or are not) modified can only occur in particular linear positions given the overall rules managing major constituent order. On the other hand, the EFP based analysis more directly addresses the descriptive principles from the literature. It directly states that a constituent must appear in a relationship with another constituent which is left-peripheral rather than relying on a side-effect of the interaction of phrase structure rules and modification. Qualitatively, the EFP might be preferred for this reason alone even if it incurs the engineering cost of extra types.

Another important consideration is how the two analyses extend to new data. Are adverbs the only lexical elements which can precede the predicate complex? How would each of these analyses fare in the face of new data? In fact, it is very common in the Lushootseed corpus that connective elements which tie sentences together in narratives precede the sentential adverbs (49). In (49a), the sentential connective *gʷəl* precedes a sentence which consists of just a sentential adverb and a verb. In (49b), the sentential connective appears to operate outside of the domain in which sentence-second position is calculated. I say this because strictly-speaking, the subject-marker *čəxʷ* appears in third position. I note that the quotative marker *kʷəł* is listed in Hess (1995, 88) as a second position particle but *yaw* is not. That leads to an analysis where the sentential connective is outside the domain of sentence-second calculation, the sentential adverb

yaw is first within the domain, and the subject-marker and quotative particle are in sentence-second position.

- (49) a. *g^wəl tuč^w ?əshudtubəx^w*
g^wəl tuč^w ?as-hud-tx^w-b=ax^w?
 CONJ just STAT=burn-ECS-PASS=NOW
 ‘They just build a fire.’ Basket Ogress (ML)
- b. *g^wəl yaw čəx^w k^wət łupig^wəd čx^wa łuhəli?*
g^wəl yaw čəx^w k^wət łu=pig^wəd čx^wa łu=həli?
 CONJ only.if 2SG.SUB QTV IRR=sing.power.song 2SG.COORD IRR=alive
 ‘And only if you put on a spirit power ceremony will you recover.’
 Crow is Sick (first telling, ML)

Under such an analysis of this data, the MODIFIED-based implementation handles the new data from (49) with little further work. The only thing required is a new rule to handle the construction: something which attaches sentential conjunctions to the left of an S-type complement. Because the analysis based on MODIFIED does not automatically percolate up to new rule types, examples like (49b) are automatically ruled in as soon as a rule for the sentential conjunction is implemented.

On the other hand, the EFP, as implemented in (44), takes the form of a constraint on *basic-binary-phrase*. To extend the EFP-based analysis to cover the new data in 49, the scope of this constraint must be reduced. One option would be to create a new subtype of *basic-binary-phrase* from which EFP-domain rule should inherit, and then to allow other phrase-structure rules to leave the PERIPH features unconstrained. This isn’t as inelegant as it may at first sound, under the consideration that the sentential connectives, by virtue of their discourse functionality, may belong in a different rule hierarchy than the one which handles sentential syntax.

A summary of the pros and cons of the two approaches is as follows. The pros of Edge Feature Principle-based approach are that it is transparent and descriptively motivated. What I

mean by transparent is that although the notion of “sentence second” is not directly expressible in a grammar built only on the recursive application of binary rules, the *L-PERIPH* feature refers to the left edge of the sentence and is mnemonically named for that function. Insofar as *L-PERIPH* maintains a pointer to the left edge of the sentence¹² it provides something which is as close to a “sentence-second” feature as possible, given the hierarchical binary rules employed in this type of grammar. The cons of this approach are (1) that independently of its reference to the left edge, the analysis still requires reference to the word/phrase opposition and (2) that the analysis doesn’t easily extend to cover the sentential conjunctions which appear in the corpus.

The pros of the *MODIFIED* analysis are (1) that it captures the relevant data within a single feature (because the lexical phrasal distinction is mirrored by whether or not a word has been modified), (2) that it extends easily to cover the sentential conjunction data within the corpus and (3) that it remains consistent with a design pattern already at play within *DELPH-IN* grammars of using the *xmod* hierarchy to constrain rule interactions in word-order analyses. The cons of this approach lie in opposition to the principal advantage of the *EFP* analysis: the machinery being used to compute second position is not named in such a way as to make it apparent exactly what work it is doing and, similarly, the usage of *MODIFIED* to mirror the word-phrase opposition hides the crucial observation which motivates the analysis and fails to connect with the simple description of the pattern to be modeled: the seeking of sentence second position by these elements.

In my current grammar system, I have moved from the *EFP*-based approach to the *MODIFIED*-based approach. The type definitions in (50) present the final versions of the constraints necessary to capture the correct placement of subject-marking auxiliaries within the system of word-order rules which already constrains word order to *VSO*.¹³

¹²Although, as seen from the discussion above, it actually maintains a pointer to the leftmost *phrase* at any level of hierarchy.

¹³There is actually a bit more to the story when it comes to these examples with sentential connectives. One of the results of the error analysis exercise described in Chapter 9 was that I ran across sentences such as (i):

- (50) `mod-head-phrase := basic-head-mod-phrase-simple & head-final &`
`[C-CONT.HOOK #h,`
`HEAD-DTR.SYNSEM [LIGHT +,`
`LOCAL.CONT.HOOK #h]].`
- `comp-aux-phrase := comp-aux-phrase-super & marker-final-phrase &`
`[MARKER-DTR.SYNSEM.MODIFIED notmod,`
`NON-MARKER-DTR.SYNSEM.MODIFIED notmod].`
- `aux-comp-phrase := comp-aux-phrase-super & marker-initial-phrase &`
`[MARKER-DTR.SYNSEM.MODIFIED hasmod].`

6.1.4 Testing and summary

The overall approach to Lushootseed word order was to collect a testsuite of sentences which represent the available word orders along with representative counterexamples for patterns which should be ruled out. These items were put into a TSDB profile to allow iterative development against the testsuite and were incorporated into the Grammar Matrix regression testing system once development was complete. The testsuite which is currently used in the regression testing system is one that was extended to cover sentences with nominal predicates. The current grammar covers the testsuite with zero overgeneration and zero ambiguity. The full results are presented in Table 6.3

In sum, after considering the built-in options for word order within the customization system

-
- (i) *hay čəd ʔuyəcəbtubicidəx^w dəg^wi siʔab dsyaʔya?*
hay čəd ʔu=yəc-b-tx^w-bucid=ax^w dəg^wi siʔab d-syaʔya?
 SCONJ 1SG.SBJ IRR=report-MD-ECS-2SG.OBJ=now 2SG noble 1SG.POSS-friend

‘And so, I am going to tell it to you, my respected friend.’

Mink and Tutyika (ES)

This example suggests that the treatment of sentential connectives as outside of the domain of sentence-second position was not correct. In the light of (50), either *yaw* in (49) is a sentence-second particle like *čəx^w* and *k^wəʔ*, or else the sentential connective *hay* has a different syntax than *g^wəʔ*.

I have added this information as a footnote to this section rather than updating the argumentation, because the reasoning as presented is what led to the selection of the *xmod*-based analysis in the grammar system which is being exposed in this document. Nevertheless, given this example, it is likely that I will return to this decision in future grammar development work.

VSOv2 Overgeneration Profile							
Aggregate	total items	negative items	word string	lexical items	distinct analyses	total results	overall coverage
	#	#	ϕ	ϕ	ϕ	#	%
$5 \leq i\text{-length} < 10$	16	12	5.00	5.00	0.00	0	0.0
$0 \leq i\text{-length} < 5$	40	24	3.00	3.00	0.00	0	0.0
Total	56	36	3.67	3.67	0.00	0	0.0

(generated by [incr tsdb()] at 1-jun-2019 (15:21 h))

VSOv2 Coverage Profile							
Aggregate	total items	positive items	word string	lexical items	distinct analyses	total results	overall coverage
	#	#	ϕ	ϕ	ϕ	#	%
$5 \leq i\text{-length} < 10$	16	4	5.00	5.00	1.00	4	100.0
$0 \leq i\text{-length} < 5$	40	16	2.75	2.75	1.00	16	100.0
Total	56	20	3.20	3.20	1.00	20	100.0

(generated by [incr tsdb()] at 1-jun-2019 (15:19 h))

Table 6.3: The results against the extended testsuite. The additional items provide coverage for sentences with nouns as main predicates.

I found that the data from Lushootseed wouldn't fit exactly into any of the preexisting buckets, but would require a new option which would mostly be able to pull from the existing VSO option, augmenting the grammar to include a class of auxiliaries which must appear in sentence-second position. Because Wackernagel's position is defined in terms of linear order, not constituency, defining the set of phrase structure rules within a constituency-based hierarchical grammar requires some level of abstraction. After considering the literature on edge marking within HPSG and taking stock of the built-in capabilities of the Grammar Matrix core-grammar, I looked at two options: one based on `EDGE` feature percolation using an implementation of Miller's EFP and the other using the `MODIFIED` feature and the `xmod` hierarchy as a proxy for lexical vs phrasal elements. I chose the latter for implementation after a qualitative analysis, taking into account the transparency of description as well as the engineering complexity of the two options in addition to how each would scale to new data already on the horizon.

	-ICS	Others
1st	-s -ubuł	-bš -buł
2nd	-sid -ubułəd	-bucid -bułəd
Reflexive	-sut	-but

Table 6.4: Lushootseed object markers

6.2 *A morphological example: object markers*

The previous section outlined the construction of a new word-order option. This section presents a morphological example: the implementation of Lushootseed object markers. The previous section attempted to shed some light on how lexical types and phrasal rules interact to license sentential constituents. In this section, I will present some of the process of word building and the morphological machinery which comes built into the customization system.

In this section, I will do my best to step gently around the topic of transitivity morphology. The object markers that I wish to discuss here operate on stems which are required to have already gone through transitivity rules. However, since the representation of Lushootseed valence-increasing morphology is the main topic of exploration in the next Chapter, here I will attempt to present Lushootseed object marking without overshadowing the upcoming content.

6.2.1 *Lushootseed Object Marking*

As discussed in §2.4.3, Lushootseed verbs are built via derivational and inflectional processes applied to intransitive roots. When a transitive stem has a first- or second-person object, this is marked on the verb according to the two allomorphic paradigms presented in Table 6.4. These paradigms are built up from four underlying morphemes: singular (-s/-bš), plural (-*(u)*buł), second person (-*(i)*d) and reflexive (-*sut*/-*but*) (Hess, 1973). As noted in the previous section, the object markers will form a part of any implementation of the rule of one nominal interpretation (39, 40) which describes the linking patterns for transitive predicates which bear a single NP argument.

Data

As in the previous section, the first step in test-driven implementation is to curate a testsuite of items which illustrate the phenomenon at hand. The testsuite I collected, presented in Table 6.5, provides concrete examples of a few broad generalizations about the data: object markers cannot attach to stems without transitivizers (0)–(7); object markers for the -ics paradigm must appear after an -ics marker (36)–(40) but may not appear after other transitivizers (32)–(36), and the converse of this is true for the elsewhere forms (8)–(15); object markers may pair with subject markers or a direct noun phrase (16)–(23), (40)–(47), but not both (24)–(27), (48)–(51); predicates generally cannot allow two direct noun phrase arguments; the negative examples in (28)–(31), (52)–(55) test that interaction with the object marker does not cause a bug in the grammar which would license this.

The testsuite as presented in Table 6.5 illustrates the surface properties of the object marker data, but it does not capture the desired constraints on semantic representations. For this, I will have to take advantage of the Grammar Matrix regression testing capabilities (described in §4.6). The methodology for implementation will be to inspect the output of the grammar by hand to verify the correct semantics, and to create a regression test based on the grammar profile upon completion of the work.

Implementation

As before, I start by seeing how far I can go using only the built-in capabilities of the customization system. In this instance, the built-in morphology system of Goodman (2013) facilitated the definition of the phonological values and morphotactics markers with no problems. All of this was accomplished through the choices file interface: I defined a position class to hold the object markers and populated the position class with 10 lexical rule types, in direct correspondence with the cells of Table 6.4. Recall that the input forms to the syntactic grammar are the underlying, or canonical, forms output by the morphophonological analyzer of Chapter 5 or the second line of Beck and Hess 2014, 2015. Each of these lexical rule types bears further choices

Number	Well-formed	Item	Number	Well-formed	Item
0	F	ʔuχ ^w -bš	28	F	ʔuχ ^w -c-bš ti luχ' ti čačas
1	F	ʔuχ ^w -buł	29	F	ʔuχ ^w -c-buł ti luχ' ti čačas
2	F	ʔuχ ^w -bucid	30	F	ʔuχ ^w -c-bucid ti luχ' ti čačas
3	F	ʔuχ ^w -bułəd	31	F	ʔuχ ^w -c-bułəd ti luχ' ti čačas
4	F	ʔuχ ^w -s	32	F	jiq-t-bš
5	F	ʔuχ ^w -sid	33	F	jiq-t-buł
6	F	ʔuχ ^w -ubuł	34	F	jiq-t-bucid
7	F	ʔuχ ^w -ubułəd	35	F	jiq-t-bułəd
8	T	ʔuχ ^w -c-bš	36	T	jiq-t-s
9	T	ʔuχ ^w -c-buł	37	T	jiq-t-ubuł
10	T	ʔuχ ^w -c-bucid	38	T	jiq-t-sid
11	T	ʔuχ ^w -c-bułəd	39	T	jiq-t-ubułəd
12	F	ʔuχ ^w -c-s	40	T	jiq-t-s ti čačas
13	F	ʔuχ ^w -c-sid	41	T	jiq-t-ubuł ti čačas
14	F	ʔuχ ^w -c-ubuł	42	T	jiq-t-sid ti čačas
15	F	ʔuχ ^w -c-ubułəd	43	T	jiq-t-ubułəd ti čačas
16	T	ʔuχ ^w -c-bš ti čačas	44	T	jiq-t-s čəx ^w
17	T	ʔuχ ^w -c-buł ti čačas	45	T	jiq-t-ubuł čəx ^w
18	T	ʔuχ ^w -c-bucid ti čačas	46	T	jiq-t-sid čəd
19	T	ʔuχ ^w -c-bułəd ti čačas	47	T	jiq-t-ubułəd čəd
20	T	ʔuχ ^w -c-bš čəx ^w	48	F	jiq-t-s čəx ^w ti čačas
21	T	ʔuχ ^w -c-buł čəx ^w	49	F	jiq-t-ubuł čəx ^w ti čačas
22	T	ʔuχ ^w -c-bucid čəd	50	F	jiq-t-sid čəd ti čačas
23	T	ʔuχ ^w -c-bułəd čəd	51	F	jiq-t-ubułəd čəd ti čačas
24	F	ʔuχ ^w -c-bš čəx ^w ti čačas	52	F	jiq-t-s ti luχ' ti čačas
25	F	ʔuχ ^w -c-buł čəx ^w ti čačas	53	F	jiq-t-ubuł ti luχ' ti čačas
26	F	ʔuχ ^w -c-bucid čəd ti čačas	54	F	jiq-t-sid ti luχ' ti čačas
27	F	ʔuχ ^w -c-bułəd čəd ti čačas	55	F	jiq-t-ubułəd ti luχ' ti čačas

Table 6.5: The development testsuite for Lushootseed object markers: These examples place the object markers of Table 6.4 across two classes of verb (*ʔuχ^w* “go”, an agent-oriented verb and *jiq*, an experiencer-oriented verb). These items were generated by taking the examples in (Hess, 1973) and creating full paradigms based on the discussion.

which constrain person and number features on the semantic index of the syntactic object. Each further bears two choices corresponding to a single lexical rule instance. One, *-inflecting*, which indicates that this rule type requires the addition of phonological material and the second, *-orth*, providing the specific value of that material. The morphotactic system also allows for the definition of *requires* and *forbids* dependencies which hold between lexical rule types, so I used this functionality to implement the *ics/elsewhere* alternation. A sample of these choices is presented in (51).

(51) verb-pc3_name=object-marking
 verb-pc3_order=suffix
 verb-pc3_lrt4_name=2plo-nonics
 verb-pc3_lrt4_feat1_name=person
 verb-pc3_lrt4_feat1_value=2nd
 verb-pc3_lrt4_feat1_head=obj
 verb-pc3_lrt4_feat2_name=number
 verb-pc3_lrt4_feat2_value=pl
 verb-pc3_lrt4_feat2_head=obj
 verb-pc3_lrt4_lril_inflecting=yes
 verb-pc3_lrt4_lril_orth=-ʔəbud
 verb-pc3_lrt4_forbid1_others=verb-pc2_lrt2
 verb-pc3_lrt5_name=1sgo-ics
 verb-pc3_lrt5_feat1_name=person
 verb-pc3_lrt5_feat1_value=1st
 verb-pc3_lrt5_feat1_head=obj
 verb-pc3_lrt5_feat2_name=number
 verb-pc3_lrt5_feat2_value=sg
 verb-pc3_lrt5_feat2_head=obj
 verb-pc3_lrt5_lril_inflecting=yes
 verb-pc3_lrt5_lril_orth=-s
 verb-pc3_lrt5_require1_others=verb-pc2_lrt2

Given this specification in the choices file, the built-in morphology system from the customization system produces a grammar with lexical rule types defined in TDL to constrain the person and number features on the syntactic object and to enforce the morphotactic constraints specified. A sample of the resulting TDL is shown in (52).

```
(52) 2plo-nonics-lex-rule := object-marking-lex-rule-super &
      [ DTR [ INFLECTED.ICS-FLAG na,
              ARG-ST.REST.FIRST.LOCAL.CONT.HOOK.INDEX.PNG [ PER 2nd,
                                                            NUM pl ] ] ].

lsgo-ics-lex-rule := object-marking-lex-rule-super &
      [ DTR [ INFLECTED.ICS-FLAG +,
              ARG-ST.REST.FIRST.LOCAL.CONT.HOOK.INDEX.PNG [ PER 1st,
                                                            NUM sg ] ] ].
```

What isn't handled by the built-in system is the implementation of the one nominal interpretation rule — that is, the syntactic constraints which are associated with this morphology. Recall that the single direct argument NP of a transitivized verb is interpreted by default as the syntactic object. Object markers can only occur on verbs which have undergone some sort of valence increasing morphology. Transitivized verbs and valence-increasing morphology is discussed in more detail in Chapter 7, but in order to show how object marking affixes were implemented, I'll say a little bit here about how the default interpretation of NP as object was implemented.

The crux of the implementation was a particular use of the feature `OPT`. `OPT`, and its associated syntax, is provided by the customization system via the argument optionality library (Saleem and Bender, 2010). A user of the argument optionality library can indicate, for subject and objects independently, whether or not argument dropping can occur and if it can, in which contexts. Users of the library can constrain argument dropping by verb-type and can indicate whether or not a marker is required on the verb (in cases of dropped arguments or in cases of overt arguments).

The choices presented in (53) are the values I have set for the Lushootseed grammar. Here is how those choices relate to the facts of Lushootseed: First, both subjects and objects may be dropped for any verb (`subj-drop-all`, `obj-drop-all`). Second, in Lushootseed, there is no marker which appears on the verb to indicate a dropped argument, nor is there a marker to indicate an overt argument (`-drop-not`, `-no-drop-not`).

- (53) section=arg-opt
 subj-drop=subj-drop-all
 subj-mark-drop=subj-mark-drop-not
 subj-mark-no-drop=subj-mark-no-drop-not
 subj-con=subj-con-always
 obj-drop=obj-drop-all
 obj-mark-drop=obj-mark-drop-not
 obj-mark-no-drop=obj-mark-no-drop-not

The library implements the actual argument dropping by providing a unary rule which removes a specified item from the relevant valence list. This unary rule of optional argument realization is constrained to require that `OPT` is set to `+`. Rules which realize the argument overtly require `OPT -`. In this way, setting `OPT +` on a particular item prevents a valence dependency from being realized through the binary phrase structure rules — only the unary *opt* rules may apply to such a valence item. The library's normal usage, then, creates verb classes which provide constraints on their arguments' `OPT` value(s). Lexical items start out `OPT bool`, so valence fillers are compatible with any constraints provided by heads or rules. Because the Lushootseed choices above allow dropping in all contexts, the verbs themselves, out-of-the-box, will not provide any further constraints. The library is still useful, however, because it sets up the infrastructure for argument dropping in a standardized way. The unadorned output of the library, when input the choices in (53), provides a grammar which allows argument dropping in all contexts, without requiring morphological markers. This is the basis upon which I implemented further constraints.

Recall that the rule of one nominal interpretation applies to two contexts: first, it says that the direct NP argument must be interpreted as object when such an interpretation is available, but secondarily, it still allows the direct NP to be interpreted as subject otherwise. This latter context is presented by intransitive verbs and transitive verbs which bear object markers. The first half of the rule of one nominal interpretation can be implemented by having the transitivity lexical rules set `OPT +` on the subject (see Chapter 7 for more on transitivity lexical rules). This blocks the application of any *head-subject-rule* which subcategorizes a full NP (because such a rule requires `OPT -` on the NP). For the second part of the rule of interpretation, things become

more interesting.

One of the contexts where an NP may be interpreted as a subject is when an object-marking lexical rule has already satisfied the object-requirement. Thus, the object-marking lexical rule will need to undo the `OPT +` constraint created by the transitivizing lexical rule, as any further NP which appears should be treated as a subject. I implemented this via the rule type shown in (54). This rule copies up the `LOCAL` and `NONLOCAL` features of the subject, but clears the `OPT` value (resetting `OPT bool`), allowing both possibilities that a *head-subject-rule* will realize an NP direct argument or the *head-opt-subject* unary rule will fill the dependency with a dropped argument.

```
(54)  change-opt-on-subj-lex-rule := lex-rule &
      [ SYNSEM.LOCAL.CAT.VAL.SUBJ < #booloptsubj & [ LOCAL #local,
                                                    NON-LOCAL #nonlocal,
                                                    OPT-CS #cs,
                                                    OPT bool ] >,
      ARG-ST.FIRST #booloptsubj,
      DTR [ SYNSEM.LOCAL.CAT.VAL.SUBJ < #optsubj & [ LOCAL #local,
                                                    OPT-CS #cs
                                                    NON-LOCAL #nonlocal ] >,
      ARG-ST.FIRST #optsubj ] ].
```

The last step is to ensure that this constraint is applied to all of the object-marking lexical rules which are output by the morphological system of the customization system. For this, I added a new file to the customization system script which manages Lushootseed object markers. This file adds to the grammar both the *change-opt-on-subj-lex-rule* type presented in (54) and a supertype which pulls in the former and insures that the `XARG` linking is correct (55).

```
(55) mylang.add(
      ""object-marking-lex-rule-super := infl-lex-rule & val-change-only
      -lex-rule &
      intransitive-lex-item & change-opt-on-subj-lex-rule &
      [ INFLECTED #infl,
        SYNSEM.LOCAL [ CAT [ HEAD #h,
                          VAL [ SPR < >,
                                SPEC < >,
                                SUBJ < [ LOCAL [ CAT.HEAD noun,
                                                CONT.HOOK.INDEX #i ] ] >,
                                COMPS < > ] ],
                          CONT #cont & [ HOOK.XARG #i ] ],

      DTR transitivizing-lex-rule-super &
      [ INFLECTED #infl,
        SYNSEM.LOCAL [ CAT.HEAD #h,
                      CONT #cont ] ] ].""
    )
```

Finally, because the morphology library code runs before the Lushootseed additions and because it assigns default lexical rule supertypes to rules which don't already have them set, a short circuit was added for particular lexical rule types which will be handled by the Lushootseed code (56). This short circuit (and the type addition above) relies on a naming convention of the morphotactics library: for each position class defined, a rule type is created with that position class' suffixed by `-lex-rule-super`. This strategy is fragile, but effective. Thus, the choices file defines a position class `object-marking`, the morphology code creates the individual lexical rule types and handles the PNG constraints requested, as well as the morphotactic constraints. The morphotactic constraints are handled by a supertype with the name `object-marking-lex-rule-super`, which the morphotactics code is told to skip when assigning semantic supertypes, and the Lushootseed `object-marking` code adds the remaining constraints to the rule.

```
(56) set_own_supertypes = ["transitivizing-lex-rule",
                          "redup-lex-rule",
                          "object-marking-lex-rule",
                          "passive-lex-rule",
                          "enclitics-lex-rule",
                          "poss-suff-lex-rule",
                          "poss-pref-lex-rule",
                          "yid-pc-lex-rule"]

if (pc.is_lex_rule and
    not pc.identifier()[:-6] in set_own_supertypes):
    if not any(st in LEX_RULE_SUPERTYPES for st in x.supertypes):
        if pc.has_incorporated_stems():
            x.supertypes.add('add-only-rule')
            x.supertypes.add('adj_incorporation-lex-rule')
        else:
            x.supertypes.add('add-only-no-ccont-rule')
```

With these modifications in place, the grammar generation system outputs a grammar which captures the testsuite in Table 6.5, rejecting the negative examples and providing just a single parse for each of the positive ones. The semantic linking was verified by hand. After this, the grammar profile and the associated testsuite were added to the regression testing system in order to ensure that future modifications don't erode this work.

6.3 Conclusion

This chapter has presented two examples of the Grammar Engineering process within a meta-grammar development environment. The first example highlighted the process of consulting the literature and comparing to previous implementations, followed by refining and making implementation decisions (which usually involve tradeoffs) for a particular phenomenon of interest. The second example illustrated how choices files and imperative program code interact in customizing a grammar. It also illustrated some aspects of the morphological and morphotactic system which will provide background for the further discussion of morphology that takes place in Chapter 7.

Chapter 7

CASE-STUDY I: VALENCE-INCREASING MORPHOLOGY

Lushootseed verbs present a rich system of valence-increasing morphology. As discussed in Chapter 2, Lushootseed is like other Salishan languages in that (nearly¹) all verbs are derived from a collection of monovalent elements, referred to in the literature alternately as either radicals or roots. Lushootseed's collection of valence-increasing morphology builds multi-argument verbs from these monovalent bases. This chapter explores the syntax and semantics of these affixes and how they may be represented in a DELPH-IN implemented grammar.

This case-study is divided into four sections: Data, Theoretical and Practical Background, Analysis, Results. The data section offers a presentation of the Lushootseed data to be modeled. I draw heavily on the work of two authors: Thom Hess, whose pedagogical work provides a wonderful clarity of data in its focus on the surface syntax and semantics which a student of Lushootseed would require in order to interact with the language (Hess, 1993, 1995); and David Beck, whose writing elegantly characterizes the data under the analytical first principles of the linguist (Beck, 2009, n.d., *inter alia*). After the presentation of data, I review the semantic literature on event-based semantics. This literature is relevant because valence-increasing affixes must be tightly integrated with semantic structures, since what they do can be seen as specifying (or possibly adding new) roles or role players to the predications of the verbs they attach to. Event-based semantics, especially the “subatomic semantics” of Parsons, is useful in framing the technical requirements of such items (Parsons, 1980, 1990, 1995). Because Parsons' work comments on and extends the earlier work of Davidson, I begin at the beginning, so to speak, reviewing the arguments and data in Davidson's foundational paper (1967). This is followed by a review of Parsons. After that, I present some of the technical arguments and refinements

¹Beck (2009) enumerates 17 underived bivalent roots.

which Dowty’s (1989; 1991) writing brings to the topic. The MRS literature builds on the event-based semantics tradition, and RMRS, specifically, implements, in a certain way, the “sub-atomic” semantics of Parsons (Copestake et al., 2005; Copestake, 2006, 2007a,b, 2009). So after reviewing the theoretical literature, I review MRS as described in several variants and as deployed in the Grammar Matrix project and the ERG (Flickinger, 2000).

After this background, I show my own work in the Analysis section, starting with an overview and diving down into the nuts-and-bolts of the types and rules which implement my approach to the data described in the first section. In the final section, I conclude by reviewing the changes to the system and the tradeoffs between the options under consideration. I offer some discussion of my own solution and how it fits into the theoretical background discussed. I give an overview of some results of the implementation, the evaluation against the test suite, noting a particular prediction of ambiguity which was turned up by the grammar but which is not addressed in the source literature (§7.4).

Seen from a high-level, this case-study provides a weaving together of the threads which form the motivation for this project: faithful discussion and representation of Lushootseed structures based on a careful review of linguistic considerations which yields an implemented apparatus that is open and extensible and which in turn yields results that improve the description of Lushootseed.

7.1 Data

In this first part of this case-study I provide a review and classification of Lushootseed verbs and valence-increasing morphology, culminating in a test-suite which forms the basis of testing and development in subsequent sections.

7.1.1 Verbs

Lushootseed valence-increasing morphology attaches only to verbal elements. Indeed, when arguing for the existence of a noun-verb distinction in Lushootseed, the attachment of valence-

increasing morphology is cited directly in Hess and van Eijk (1985) (and obliquely in van Eijk and Hess (1986)) as one of the characteristics which indicate that a particular stem must be verbal. Because this chapter treats the syntax and semantics of these affixes, and because their syntax and semantics interacts with that of the bases to which they attach, this section begins with a review of Lushootseed verb classes.

The introductory grammar which precedes the texts of *Lushootseed Reader Volume I* (Hess, 1995, henceforth: LR1) provides a convenient introductory framework for Lushootseed verb classes. However, due to the pedagogical nature of the material, Hess' claims and generalizations about verb structures are indirect. Hess presents paradigms of examples and asks the reader whether particular arguments are functioning as “agent (the doer)” or “patient (the one to whom some act is done)”, eliciting the pattern from the reader without explicitly staking a claim for a particular theoretical approach. Naturally, such a work emphasizes the understanding of fully inflected forms over specific rules that might predict morphological patterns. Although Hess consistently refers to stems, not roots, in his verbal analyses, if I strip away the morphology and examine what's left, LR1 appears to contrast the underlying verbal roots as having agent-orientation or experiencer-orientation. Stems built from these roots without further suffixation never accept patient arguments. Patient-orientation is always derived using what Hess calls a “patient suffix”: the collection of valence-increasing morphology which is under study in this chapter.²

Despite the indirect presentation of linguistic generalizations, the first half of LR1's pedagogical grammar (“Unit 1: identifying semantic roles in the main clause”), presents the Lushootseed valence-increasing affixes in paradigmatic groups, illustrated against a small vocabulary of verbs, and culminates in a summary of the output of these morphological processes into five stem classes.³ Because of its paradigmatic presentation, and its constrained vocabulary, the 237

²Note that a “patient suffix” is distinct from the object markers which are presented and analyzed in Chapter 6.

³The five stem classes of LR1 are:

- Agent Oriented
- Middle Voice

sentences of this Unit form (most of) the test suite for this case-study.⁴

In the next section, I review the affixes themselves for their morphological, syntactic and semantic properties. From the standpoint of preparing for that section, the takeaway here is that Lushootseed verbal roots are generally monovalent and that both unaccusative as well as unergative forms exist. Hess suggests that these underlying orientations are implicated in the predication of morphological patterns, but in LR1 the entire system is not made explicit for the pedagogical audience. Hess speaks more directly about his categorization of Lushootseed verbs and verbal morphology when addressing the linguistic audience in Hess (1993). In that work, he argues for primary attention to be paid to the derived stem (as opposed to the root) when building classifications:

The root is easily discernible in the vast majority of Lushootseed verbs; however, it is the stem and not the root that is the basic descriptive unit of that verb. The descriptive primacy of the stem has gone unnoticed for many years due to the salience of the root (almost always CVC) which attracted (or distracted) the researchers' attention and because several stem classes turned out to have identical shapes in most utterances effectively camouflaging significant differences among them. Then, too, in many verbs the root and stem are in fact coterminous as was once believed true of all Lushootseed verbs. (Hess, 1993, p. 113)

I will return to this notion of descriptive primacy and the prediction of verbal morphology from roots below when I discuss my own work. Also of note in understanding Hess' point of

-
- Patient/Goal Oriented
 - Passive (noted as a subtype of Patient/Goal oriented)
 - Experiencer Oriented

Verbs without valence-increasing affixes always appear in the Agent Oriented or Experiencer Oriented classes.

⁴Another value of using the sentences from Unit 1 of LR1 as the case-study is the ancillary interacting phenomena which occur. In addition to the subject marking auxiliaries discussed in Chapter 6, there are question marking auxiliaries and passivization. Question marking auxiliaries require sentence-second position, so their analysis uses the same positioning schemata described in Chapter 6. Passivization occurs in complementary distribution with the object marking, so it uses the same position class to enforce this distributional pattern. However, the argument structure and syntactic constraints are not the same as the object markers.

view is that for him, the term “transitive” refers to a syntactic property, not a semantic one, as revealed by the following passage in which Hess rejects the Salishanist tradition of referring to valence-increasing suffixes as “transitivizing”.

The preceding brief remarks present a different perspective from what is typical in Salish studies. The class of patient suffixes has in all previous descriptions of Lushootseed been called transitive (undoubtedly due to the influence of the English glosses); and their cognates are labelled transitive suffixes in all grammars of other Salish languages (we have read). However, in Lushootseed, at least, these suffixes do not mark transitivity. With only one direct argument in the predicate — whether it marks an agent or a patient, it makes little sense to talk about transitivity. Instead, these suffixes contribute to identify the semantic role of that argument. (Hess, 1993, p. 117)

To summarize, from Hess, and specifically from the first half of the LR1 grammar, I have taken the main collection of examples which serves as a test-suite in developing my implementation. Furthermore, I take away the fact that Lushootseed roots can be thought of as contrasting an agent vs experiencer “orientation”, but note that this underlying orientation doesn’t actually provide a full prediction of which valence-increasing affixes combine with each root. For that, I would need to collect Lushootseed verbs into stem classes. Hess provides an outline of such classes in Hess (1993), but I have not fully implemented this approach, because that paper only has a handful of examples. My approach to stem classes is given below in §7.3.2. Hess’ approach to transitivity is also not one that I wholly adopt, in that I have had some success with the approach to word-order outlined in Chapter 6.

7.1.2 Valence increasing morphology

I turn to David Beck’s description of the valence changing affixes in order to introduce them here (Beck, 2009, n.d.). In this section, I step through the affix types, following Beck’s taxonomy and typology.

Beck's classification divides Lushootseed valence-increasing morphology broadly into causatives and applicatives. The term "causative", in this context, as I understand it, merits a caveat because of the fact that later in the chapter I will be discussing semantic role theory wherein the term CAUSER may imply a special, technical meaning. Here, a "causative" simply refers to the morphosyntactic property of adding a syntactic subject. It is not necessarily the case that these subjects are CAUSERS in the sense indicated by semantic role theory — although more will be said on this below. The morphosyntax of causatives can be contrasted with that of the applicatives, which, on the other hand, retain the subject of their base and add a new role which is linked to either a syntactic object, or an oblique — this latter contrast forming a sub-classification among the applicatives. The classification of causative and applicative in this subsection is strictly based upon the manipulation of grammatical functions (and their linking). To put things bluntly, "causative" in opposition to "applicative" will refer to valence-increasing morphology which is subject-adding as opposed to complement- (or object-) adding. This usage of "causative" is distinct from the question of whether or not the added subject is a semantic CAUSER.

In this case-study I assume that it's completely possible to abstract lexical semantic properties of words away via symbolization, leaving only the structural semantic properties. So Beck's classification of these affixes strictly by morphosyntax forms an ontology which is congruent with my objective. Nevertheless, in the following descriptions I report a summary of Beck's notes on the lexical semantics of these elements in the hope that the descriptive content of these summaries may help the reader to connect the form to the gloss in the analysis below.

7.1.3 *Causatives*

Lushootseed causative markers are subdivided according to the two types of argument structure they produce: transitive and intransitive. While the transitive causatives still obey the rule of one-nominal interpretation which prevents the realization of more than one full NP argument in a clause, they produce otherwise a straightforward transitive, a verb which subcategorizes for

a subject and direct object.⁵ The intransitive causatives, on the other hand, demote the former subject not to direct object, but to oblique. These forms stand out in Lushootseed morphosyntax because when they have a single NP argument, this argument is interpreted as subject — they accept no direct object NP. Below I provide examples of each of these subtypes.

I will refer throughout this chapter, then to "transitives" in opposition to "bivalent intransitives". The bivalent intransitives are syntactically bivalent because they subcategorize for two arguments, but they are distinct from the (also bivalent) transitives in how their arguments are marked by the syntax. That said, bivalent, for me, simply indicates the syntactic arity of a verb, irrespective of the argument marking which is subcategorized for. The terms intransitive and transitive refer directly to those marking patterns.

Transitive Causatives

Lushootseed provides three suffixes which add a syntactic subject to an intransitive base, demoting the erstwhile subject to syntactic object (Beck, n.d., §2.1.12). Beck provides an overview of the semantic characteristics which distinguish these affixes, but as I mentioned above, in my treatment, these lexical-semantic differences have been abstracted away from. This abstraction in itself, however, constitutes a hypothesis that the structural facts can be accommodated without reference to lexical-semantic properties. The one reflection of these semantic characteristics which is present herein is the naming scheme for these forms: the "internal" (57), "external" (58), and "diminished control" (59) are so named by Beck in reference to their (lexical-)semantic characteristics. I also follow Beck's convention in glossing abbreviations (ICS, ECS and DC).

The transitive causatives are in complementary distribution. On any given token no more than a single transitive-causative may apply⁶ (modulo the appearance of the internal causative with the secondary suffixes, see below). Examples (57), (58) and (59) illustrate the three tran-

⁵See (39) and the surrounding discussion in Chapter 6 for the context of the rule of one-nominal interpretation and its effect on Lushootseed syntax.

⁶There are morphophonological selection criteria, however, which make reference to these affixes: allophonic variation in the object markers (discussed in Chapter 6) depends on the specific presence of the internal causative.

sitive causatives.

- (57) a. *ʔuʃiǰ* *ti* *čáčas*
 ʔu-ʃiǰ *ti* *čáčas*
 PFV-be.immersed DET child

‘The boy drowned.’

- b. *ʔuʃiǰəd*
 ʔu-ʃiǰ-t
 PFV-be.immersed-ICS

‘He immersed something.’

Hess, 1973, p. 89

- (58) a. *ʔuʔuχ^w čəd*
 ʔu-ʔuχ^w čəd
 PFV-go 1SG.SBJ

‘I went.’

(Hess 1995: 6, ex. 1)

- b. *ʔuʔuχ^wtubš* *ti* *čáčas*
 ʔu-ʔuχ^w-tx^w-bš *ti* *čáčas*
 PFV-go-ECS-1SG.OBJ DET child

‘The boy took me.’

(From Beck: based on Hess 1995: 42)

- (59) a. *ʔuk^wəłəd* *ti* *q^wu?*
 ʔu-k^wəł-t *ti* *q^wu?*
 PFV-pour-ICS DET water

‘[Someone] poured the water.’

- b. *ʔukʷətdxʷ ti qʷu?*
ʔu-kʷətdxʷ ti qʷu?
 PFV-pour-DC DET water

‘[Someone] spilled the water.’

Hess 1995

From a very high level, the lexical-semantic difference between the internal and external causative is that in the external, the causer is “less directly involved in or affected by the event than in stems formed with the internal causative” (Beck, n.d., §2.1.2.2). Then, in the third type of causative, diminished control, the added subject acts with less than full control over the outcome of the event.

Hess and Bates (1998) provide an interesting and detailed analysis of subtypes of *-txʷ* (the morpheme referred to by Beck, and here, as the external causative). In this paper, the authors work from an ergative-absolutive analysis and divide examples with *-txʷ* into subtypes: those in which the absolutive argument is a causer but not an agent (60) and those in which it is both a causer and agent (58b) and others.

- (60) *yayustxʷ tsi sʔadəy?*
yayus-txʷ tsi sʔadəy?
 work-ECS DET woman

‘She put the woman to work.’

Hess and Bates (1998)

However, as mentioned above, my approach abstracts away this semantic substructure. At the risk of getting ahead of myself in the exposition, I’ll say a bit about why I think this is feasible: Even though (Hess and Bates, 1998) identify a difference in the lexical semantics of these two examples, the structural semantics can be seen as the same. In both of these cases, the syntactic subject is associated with the argument which has been added by the morphology and the object is associated with the argument of the underlying intransitive. So (60) has an expressed argument for *yayus* (“work”) and an unexpressed argument for *ECS* while (58b) has a grammatically

marked 1SG argument for $\text{?u?u}\check{x}^w$ and an expressed argument for ECS. That these distinct semantic facts are so succinctly unified by this symbolic approach, to me, provides evidence that the approach I've taken to the syntax of these elements (and their relation to grammatical functions) is on the right track.

Intransitive Causatives

The intransitive causatives attach to an intransitive stem and yield a form where the internal subject has been demoted to an oblique phrase, and there is no direct object: any phrase which is in direct argument position must be the syntactic subject. The two causative affixes of this class are called the causative middle (61) and the causative of activity (62).

Beck describes the semantic difference between the causative middle and the other causatives as being one in which focus is placed on the contribution of the agent, or on the agent's self-interest in the result (Beck, n.d., §2.1.2.4). The first two sentences of (61) show how the same verb stem can be compatible with either the internal causative or the causative middle, illustrating additionally the two different argument structures produced by the two types of causative and their interaction with the rule of one nominal interpretation. The example in (61c) shows the oblique marking subcategorization which the causative middle demands.

- (61) a. $\text{?ug}^w\text{ə}\check{\text{c}}\text{ə}d$ *tsi* $\check{\text{c}}\text{ə}\check{\text{c}}\text{as}$
 $\text{?u-g}^w\text{ə}\check{\text{c}}\text{-t}$ *tsi* $\check{\text{c}}\text{ə}\check{\text{c}}\text{as}$
 PFV-search-ICS DET.F child
 '[Someone] looked for the girl.'
- b. $\text{?ug}^w\text{ə}\check{\text{c}}\text{ə}b$ *tsi* $\check{\text{c}}\text{ə}\check{\text{c}}\text{as}$
 $\text{?u-g}^w\text{ə}\check{\text{c}}\text{-b}$ *tsi* $\check{\text{c}}\text{ə}\check{\text{c}}\text{as}$
 PFV-search-CSMD DET.F child
 'The girl looked for [something/someone].'

c. *ʔug^wəčəb* *ti* *čəčas* *ʔə* *ti* *sqəlalitut*
ʔu-g^wəč-b *ti* *čəčas* *ʔə* *ti* *sqəlalitut*
 PFV-search-CSMD DET child OBL DET spirit.power

‘The boy quested for spirit power.’

(Hess, 1995)

The second affix in the class of bivalent intransitive-forming causatives is called the causative of activity. This affix attaches to action verbs yielding a stem with a meaning where the semantic subject engages in the activity expressed by the verb in such a way as to affect or involve the internal subject of the intransitive base (62) (Beck, n.d., §2.1.2.5).

(62) *ʔušabalik^w* *tsi* *luλ'* *ʔə* *ti* *sʔuladx^w*
ʔu-šab-alik^w *tsi* *luλ'* *ʔə* *ti* *sʔuladx^w*
 PFV-dry-ACT DET.F older.person OBL DET salmon

‘The old woman dried salmon.’

LR1 p.28

7.1.4 *Applicatives*

In this section, I continue outlining Beck’s classification of valence-increasing morphology, now turning to the applicatives. Beck’s taxonomy of applicatives includes a single, straightforward applicative alongside the group of “secondary suffixes” which create trivalent verbs and which do not occur without some support from the internal causative. In categorizing applicative affixes, Beck also presents a notable example which may pose some difficulty for the clean classification of valence-increasing affixes into pure causatives and pure applicatives: the applicative usage of the internal causative.

Allative applicative

Beck’s n.d., §2.1.2.6 description presents the allative applicative as the most straightforward of the Lushootseed applicatives. This affix adds a new argument to an intransitive base which is

usually a semantic GOAL. Hess (1973) provides examples, some of which are shown here in (63).

- (63) a. *ʔuʔuχ^w ti čáčas*
 ʔu-ʔuχ^w ti čáčas
 PFV-go DET child

‘The boy went.’

- b. *ʔuʔuχ^wc*
 ʔu-ʔuχ^w-c
 PFV-go-ALTV

‘He went after someone.’

- c. *ʔug^wədil ti čáčas*
 ʔu-g^wədil ti čáčas
 PFV-sit DET child

‘The boy sat.’

- d. *ʔug^wədils*
 ʔu-g^wədil-s
 PFV-sit-ALTV

‘He sat next to someone.’

Hess 1973

The examples in (64) make the syntax explicit. The subject of an intransitive verb, *ʔalil* (“go ashore”) is retained in its agent-like role even after the addition of the ALTV marker (*-s* is an allomorph which appears after *il*). Recall that according to the syntactic rule of one-nominal interpretation,⁷ *ti supq̄s* (“hair seal”) is the syntactic object by virtue of being the direct NP argument of a verb without an object-marker, illustrated in (64b) and (64c).

⁷Sec (40) of Chapter 6.

- (64) a. *ʔuʔalil ti luʔ*
 ʔu-ʔalil ti luʔ
 PFV-go.ashore DET old.person

‘The old man went ashore.’

- b. *ʔuʔalils ti supqs*
 ʔu-ʔalil-s ti supqs
 PFV-go.ashore-ALTV DET hair.seal

‘He went ashore after a hair seal.’

- c. *ʔuʔalils čəd ti supqs*
 ʔu-ʔalil-s čəd ti supqs
 PFV-go.ashore-ALTV 1SG.SBJ DET hair.seal

‘I went ashore after a hair seal.’

LR1

Other applicatives

There is a secondary set of applicative-like affixes which cannot appear without the support of a primary affix (which is almost always the internal causative).⁸ Hess (1993) lists three of these “secondary affixes”: *-yi-*, *-bi-* and *-di-*. Hess and Bates (2004) expand this list with *-i-*, but Beck (n.d., p. 103) notes that of these four, only *-yi-* and *-bi-* appear to have any productivity. I look first at *-yi-*.

When it comes to glossing the meaning of *-yi-*, or offering a semantic role label, Hess (1995) demurs. In contrast to his use of “agent”, “patient” and “experiencer” as labels for the semantic roles associated with other constructions, LR1 refers to the argument of *-yi-* as bearing simply

⁸ While Hess (1993) claimed that *-yi-* was universally followed by the internal causative, Beck (n.d., p. 103) presents a counterexample (and claims it is the only such in the Lushootseed Corpus) wherein a secondary suffix is followed by a different suffix. The verb *cil* “to dish out” appears with the dative applicative followed by the causative of activity (on line 76 of Martha Lamont’s second telling of “Crow is Sick”) (Hess, 1998, p. 63). This aside highlights the utility of digitized corpora — even when those corpora are relatively small.

the “-yi-” role.⁹ Beck refers to the -yi suffix as the “dative applicative” (DAT) and suggests that the semantic role it adds is a RECIPIENT or a BENEFICIARY depending on the context provided by the base.

As with the other suffixes, I follow Beck’s glossing, so I’ll refer to this morpheme as DAT in examples. A second note which bears on representation, however, is that because -yi- is always followed by ICS (modulo one *hapax legomenon*⁸), I treat the sequence *yi-t* (DAT-ICS) as a single morphological complex. The implementation details for this morpheme (sequence) are discussed in the Analysis section below (see §7.3.2). Despite the fact that I am gluing together the two parts in my implementation, the sequence will continue to appear with an internal hyphen both in this text and in the testsuite, maintaining a consistency with the representation in the source data.

Now I turn to examples. The sentence in (65) is contrasted to (40a). The direct argument of *ʔuʔuχ^wyi^d* bears the -yi- role. Although Hess’ free translation of this sentence suggests only two role players (the boy and the one who went for the boy), Beck claims that “when the morphological complex -yi-d is added to a monovalent intransitive base, the effect is an increase in valency of two” (Beck, n.d., 104). Indeed there are other examples in the corpus which make the three arguments of such a construction explicit such as (65b).

- (65) a. *ʔuʔuχ^wyi^d ti čáčas.*
 ʔu-ʔuχ^w-yi-t ti čáčas
 PFV-GO-DAT-ICS DET child

‘[Someone] went instead of the boy.’ (Someone went so he wouldn’t have to.)

(Hess, 1995, p. 34)

⁹Hess explains this incongruence as follows:

There are least five terms in linguistics used to convey the role [...] above. These are *benefactive*, *dative*, *indirect object*, *recipient*, and *second object*. None of these is fully satisfactory for Lushootseed. Therefore, in Lushootseed grammar one speaks simply of the **-yi-role**. (Hess, 1993, p. 34)

- b. *ʔuk^wədyic* *ʔə ti ʔaʔx.*
 ʔu-k^wəd-yi-t-s ʔə ti ʔaʔx
 PFV-taken-DAT-ICS-1SG.OBJ P DET platter

‘S/he took the platter from me.’

(Hess, 1995, p. 42)

Applicative ICS

Lushootseed grammar contains a form which is homophonous to the internal causative (ics) morpheme discussed above, but which patterns morphosyntactically as an applicative, adding a direct object to the verb’s subcategorization pattern rather adding a new subject and demoting the existing one. This usage of the ics morpheme appears to be restricted to its application to agent-oriented stems (see Table 23 of Beck n.d.).

Consider the examples in (66a). The first sentence shows that the base form of the stem *ʔil* (‘sing’) is built around an agent-oriented (unergative) predicate, whose single argument is the singer. In the second sentence of (66b), the singer argument is still the subject (indicated by the subject marking auxiliary *čəd*), so the additional argument added by the internal causative expresses the what the song is about. I have bracketed the direct NP argument in order to highlight the structure. Thus, the construction is technically an applicative, not a causative.

- (66) a. *huy ʔiləx^w tsiʔəʔ kaʔkaʔ*
huy ʔil=əx^w tsiʔəʔ kaʔkaʔ
 SCONJ sing=NOW DET.F CROW

‘Then, Crow sings.’

applications of morphology. This view is extremely useful in understanding the rule of one-nominal interpretation and the object marking paradigm, but one potentially questionable item which falls out of this approach is the lack of unity in describing the syntactic exponence of the ICS morpheme. Hess' model treats both applicatives and causatives alike by viewing them as patient marking suffixes. This unifies the applicative and causative uses of *-t* (ICS) by saying that in both cases, the morpheme indicates that the direct argument is patient-like — of course, saying such a thing gets us no closer than positing homophonous forms when the objective is to describe lexical rules which map from the base form (and semantic linkings therein) to the output.

Beck (2009, 542, 543) notes that the differences in these two descriptions may boil down to whether one views Lushootseed valence-increasing morphology as inflectional or derivational. On the inflectional view, the appearance of a “patient suffix” co-occurs with an overt patient-argument of what is underlyingly a transitive event structure. On the derivational approach, the bare roots describe intransitive situations and the affix derives a new structure with more participants. I find Beck's arguments regarding the lexically determined meaning of the causatively marked stems (Beck, 2009, 545) convincing enough to form the basis of my implementation (presented below). Conceptually, this approach is consistent with Dowty's (1991) predictions regarding the ordering and indexing of arguments along proto-Agent and proto-Patient dimensions (which is also discussed below).

Hess' claim that treating “patient suffixes” as the first inflectional affix (Hess, 1993, 118) leads to a salient hierarchy of stem classes is an attractive one to pursue, from a grammar engineering perspective, but I have set this idea aside for to pursue in future work. I haven't yet attempted an implementation of Hess' approach.

7.1.6 Testsuite

The full testsuite for this case-study is all of the sentences of LR1 Unit 1, along with the examples from Hess (1976) which illustrate the valence-changing affixes. There are a total of 259 sentences in the suite. The entire testsuite is listed in Appendix B. In addition to the affixes dis-

cussed above, the testsuite covers interacting phenomena such as question marking auxiliaries and passivization.

7.1.7 Summary

The first section of this case-study has presented an overview of the Lushootseed verb forms as well as the valence-increasing morphology. From a high level, the descriptive takeaway is that most verbs are underlyingly monovalent. Beck classifies Lushootseed radicals as agent or patient oriented in reference to the semantic role of the single argument of such monovalent roots. Hess, on the other hand, focuses on Lushootseed stems: treating the valence-increasing affixes as inflectional markers which may produce patient-oriented stems. The affixes themselves can be descriptively classified into causatives and applicatives. Beck's description of the lexical semantics of the three causatives is mentioned, but analytically I consider these notes to be abstracted away in my own work; I use the names which Beck has coined: the internal, external and direct causative (glossed *ics*, *ecs* and *dc*) as symbols to stand in for the lexical semantics. For applicatives there is a single "simple" applicative alongside the secondary suffixes which are used to form trivalent verbs. Finally, there is some limited usage of a form which bears the same shape as the internal causative morpheme but which creates an applicative construction. Analytically, a system such as Hess', which turns on the construction of patient-oriented stems, unifies these otherwise contrasting usages of the *ics* morpheme. However, the derivational system described by Beck provides a more complete starting point for implementation.

7.2 Theoretical and Practical Background

Having examined the data to be modeled, I now turn to the semantic literature for background on how to model it. Valence-changing morphology provides a test case for exploring the border between structural semantics and lexical semantics. This is because it must alter or interact in some way with the predicate structures indicated by the verbs it attaches to and it must do so in a way which is consistent across the lexical classes of verbs to which it may apply. The

semantic representations in DELPH-IN grammars owe an intellectual debt to the tradition of event-based semantics, so the first part of this section traces some of the history of ideas which have informed the writers of MRS semantics and its variants. Then, particular aspects of MRS itself are explored in the subsequent part. These two background discussions are followed by a reflection on the application of these ideas to the data of the Lushootseed testsuite.

7.2.1 *Event-based semantics*

Event-based semantics (Davidson, 1967, 1985; Parsons, 1980, 1990, 1995; Dowty, 1989, 1991), also called Davidsonian semantics, is a semantic theory built on the hypothesis that verbal elements do not directly talk about a particular event, but instead provide constraints on an underlying event variable, which is bound by a quantifier within a particular scope. The idea that verbal meaning should be discussed in terms of a quantified variable, rather than direct reference, is analogous to what has been standardly assumed about nominal phrases.¹⁰ With the exception of certain treatments of proper names, in most semantic theories bare nominals denote sets of individuals with particular properties. A nominal, or substantive, become referential only after quantification. The latter is typically represented in logical formulas by binding a variable within a scope and asserting the properties which characterize the set as predications of the variable within that scope. Often these properties are thought of as expressing the truth-conditions of an expression as in Tarski's program (Hodges, 2018). This idea that verbal elements should also be represented as presenting constraints which hold on a variable ranging over events is usually attributed to Davidson (1967).

In this section I recap Davidson's original argumentation, followed by a review of Parsons' (1990; 1995) system, which further decomposes structures and builds a particular theory of thematic roles. After this, I outline Dowty's (1989) view of thematic role theory and semantic analysis. Finally, I go over Dowty's hypothesis of Proto-roles from which linguistic argument structures can be predicted. This preliminary discussion provides the lead-in to the discussion

¹⁰See Geach 1962 for an overview of logical theories of meaning. Chapter 3 treats referring phrases.

of the MRS variants available within the DELPH-IN universe, which eventually leads to my proposed representation of the Lushootseed data.

Davidson

In “The Logical Form of Action Sentences”, Donald Davidson begins with consideration of the sentences in (67).

- (67) a. Jones buttered the toast in the bathroom with a knife at midnight.
 b. Jones buttered the toast.

Truth-conditionally, (67a) implies (67b). For Davidson, this implication should therefore be transparent in the semantic representations of these sentences. If (67a) is treated as a five place predicate (*buttered*(x,y,z,v,w)) then (67b) must be treated as elliptical for “Jones buttered the toast somewhere with something at some time” in order to capture the implication — because without predicate mapping rules, the predicate symbols with distinct arities are not otherwise connected to each other. However, the idea that the semantic structure of every verb in a language must carry an argument position for every type of restrictive predication¹¹ with which it may appear is unattractive because of the very large predicate signatures which result. But, alas, treating (67b) as a simple two place predicate (*buttered*(x, y)) removes the implication which should hold between (67a) and (67b), since the two predicates are now distinct symbols, with different argument signatures.

Davidson’s solution to this is to posit an underlying variable which represents an event alongside a conjunction of predicates which make reference to that variable, as in (68).¹² The repre-

¹¹The term, “restrictive predication”, as I have used it here, refers to the types of predications which provide additional constraints, or truth conditions, on the sentence. These are called restrictive because of the subset relation which they bear to sentence interpretation: adding “with a knife” to the sentence in (67b) describes a subset of the events described by (67b). The independent and restrictive nature of such predications allows them to be treated as additional conjuncts, the observation of which is a contribution of Davidson’s work.

¹²Modern Davidsonian semantics typically places the underlying event(uality) variable first, rather than last, and symbolizes it as e rather than an x (in reference to a theory of semantic variable types). In this discussion, I have maintained Davidson’s original representation convention.

sentation in (68) provides the desired implication between (67a) and (67b) because (68b) asserts the truth of a subset of the predicates asserted in (68a).

- (68) a. $(\exists x)$ [battered(Jones, the toast, x) &
 in(the bathroom, x) &
 with(a knife, x) &
 at(midnight, x)]
- b. $(\exists x)$ [battered(Jones, the toast, x)]

The next question which arises, in considering these representations, involves specifying the criteria for deciding which of a verb's arguments to include in the core representation of the predicate symbol and which to factor off as separate predicates. Davidson's original answer invokes English prepositional syntax: "in general, we conceal logical structure when we treat prepositions as integral parts of verbs; it is a merit of the present proposal that it suggests a way of treating prepositions as contributing structure" (1967, 93).

Davidson, then, argues for the explicit conjunction of any prepositional material, retaining the core arguments of a verb as integral parts of its predicate signature. While Davidson's criterion may seem overly particular to English, I understand the generally applicable linguistic hypothesis which may be drawn from this to be that languages may provide idiosyncratic means for distinguishing which elements pertain to a verb's core semantics and which provide additional conjuncts, but that there ought to be some morphosyntactic reflection of the semantic structure, should the criterion stand. In a later work, Davidson refined his criteria as below, saying that we must

...reduce the number of places of the underlying verbal predicate to the smallest number that will yield, with appropriate singular terms, a complete sentence. But do not think you have a complete sentence until you have uncovered enough structure to validate all inferences you consider due to logical form. (Davidson, 1985, 232–33)

Davidson’s ideas now form the basis of one of the main branches of the theory of meaning (Speaks, 2014). Event-based semantics is a direct influence on situation semantics (Barwise and Perry, 1983; Kratzer, 2017), which, in turn, played an influential part in the development of the original semantic framework of HPSG (Pollard and Sag, 1994, 24–24, 318). While the introduction of MRS significantly changed how many HPSG grammars deal with quantifiers and scope, Davidsonian event-based semantics still underpins the MRS approach to role-players and predications.

Parsons

Terence Parsons (1990; 1995) accepts Davidson’s arguments for decomposing verbs into a conjunction of predications, but rejects the idea that the core arguments should be included in the verb’s predicate signature. Parsons’ work on “subatomic semantics” combines thematic role theory with Davidson’s system of quantification over event variables, in that the core arguments are conjoined to the verb using thematic role predicates, as shown in (69).

- (69) a. Brutus stabbed Caesar violently in the back with a knife in the agora.
- b. $(\exists e)[\text{Stabbing}(e) \ \&$
 Agent(e , Brutus) &
 Theme(e , Caesar) &
 Violent(e) &
 In_{TARGET}(e , the back) &
 With(e , the knife) &
 In_{LOCATION}(e , the agora)]

The names of the core-argument predications are drawn from a collection of linguistic thematic roles: AGENT, THEME, etc. So, in (69), the main predicate, *Stabbing*, instead of having an arity of three (as would be the case in in Davidson’s original analysis), has arity of one, retaining only the event parameter. The predicates which connect Brutus and Caesar into the event do so

by also taking an event parameter. The role players are then added into the semantic representation via a conjunction of additional predicates which assert their thematic relations to the event in much the same way as verbal modifiers do under Davidson's original proposal.

Parsons' motivation for further decomposing predicate structure and adding additional conjuncts to represent thematic roles is connected to historical argumentation for the utility of thematic roles in semantic explication. In establishing the precedent and the utility of thematic roles in linguistic analysis, Parsons goes both ancient and contemporary, citing Pāṇini as well as the Case Grammar of Fillmore (1968). But beyond these references, Parsons adds his own argumentation for decomposition into independent semantic-role predications based on the lack of syntactic marking to distinguish core arguments from modifiers in many languages. Parsons claims that his system provides the correct application of Davidson's 1985 refined criterion for deciding when to incorporate an argument into a predicate structure and when to create an independent conjunct.

According to Parsons, Davidson's refined criteria (quoted above), when considered under the fact that for some predicates in some contexts, even the core role-players may be optional as well as the fact that natural language provides the expressive ability to describe impossible situations (Parsons, 1990, 95–99), amounts to an argument for full decomposition. For the purposes of grammar engineers, setting aside for the moment the issue of the names of the thematic roles themselves, considering only the matter of full decomposition, Parsonian representation is also motivated along practical lines. This is because many natural language processing tasks lead to scenarios in which only partial information about a structure is available — a point raised by Copestake (2006) as a key factor in the creation of RMRS (discussed below).

Taking, for the moment, the hypothesis that Parsonian decomposition is motivated by both practical and theoretical concerns, I turn to the matter of the thematic role labels themselves. This topic is treated in depth in a pair of papers by David Dowty which provide a high-level overview of thematic roles in semantic theory, looking specifically at their usage in a system like the one proposed by Parsons.

Dowty

David Dowty (1989) takes up the challenge implicit in the observation that thematic roles, as employed by contemporary syntactic explorations (Gruber, 1965; Fillmore, 1968; Jackendoff, 1972, *inter alia*), including the work of Parsons, are useful in a descriptive sense, but too vague for model-theoretic application.¹³ Dowty, like Parsons, acknowledges the intellectual heritage that the analysis and theory of thematic roles in linguistics owes to the grammar of Pāṇini, but refines this, connecting their contemporary usage in late twentieth century syntax and semantics to an MIT dissertation by Gruber (1965) as well as the Case Grammar project of Fillmore (1968). Dowty's chapter explores two approaches to a *semantically-based* theory of thematic roles: an "ordered argument" approach which defines thematic roles as collections of entailments which hold between arguments of verbs and a "neo-Davidsonian" approach along the lines specified in Parsons (1980).

Ordered-argument system An ordered-argument system is, in many respects, an attempt to maintain the traditional approach to predicate-argument structure in which the definition of a predicate includes a signature of arguments which are distinguished from each other via their order. Dowty explores how the notion of thematic roles plays out formally in such a framework.

The key idea which connects thematic roles to this traditional approach is to see thematic roles as collections of entailments which hold of arguments of particular verbs. If *x builds y* then the lexical entailments of *builds* provide particular information about *x*: that *x* acts volitionally, etc. For Dowty, the particular set of entailments about *x* which can be known from the truth of *x builds y* can be thought of as an *individual* thematic role. Such an individual thematic role, in this example, is effectively the "builder" role: it encompasses all the things entailed by being the subject of build.

But the goal of Dowty's first system is to use these individual roles — collections of entailments — as a building block for a system of thematic role types (Dowty, 1989, 77,78). To this

¹³See also Chierchia 1989.

end, he defines a thematic role *type* as a particular intersection of entailments. The idea being that the familiar linguistic thematic role types: AGENT, THEME, etc might be defined by taking the collections of entailments which are held in common for all the verbs which assign AGENT to their subject, and computing their intersection.

Any collection of entailments, however, under this system is, by definition, a thematic role type. For example, the particular intersection of entailments which hold of the subject of *build* along with those of the indirect object of *give*, would form a thematic role type — although not a very intuitive one. In order to distinguish the set of entailments which are of descriptive interest — those types which are highlighted by virtue of their commonality across the verb classes of a language — Dowty distinguishes the *L*(inguistic)-thematic role types as a subset of all thematic role types. Dowty enumerates three properties of an L-thematic role type system which he suggests should hold.

1. Completeness

Every individual thematic role contains some L-thematic role type (or as we may equivalently say, every argument position of every verb is “assigned” an L-thematic role type) (Dowty, 1989, 78)

2. Distinctness

Every argument-position of every verb is distinguished from every other argument-position of the same verb by the L-thematic role types the two argument positions are assigned. (Dowty, 1989, 78)

3. Independence

The properties in an L-thematic role type must be characterizable independently of the relations (denoted by natural language verbs) that entail them. (Dowty, 1989, 82)

With the above constraints, the ordered-argument definition of thematic role types provides both a semantic basis for the actual model-theoretic properties of thematic roles as well as their usage in systems of linguistic description. Such a system can form the basis of hypothesizing about, for example, the semantic aspects of obliqueness hierarchies. In contrast to the ordered-argument system, Dowty presents a second system under which the above criteria may also be examined.

Neo-Davidsonian system Dowty (1989, 84) credits Parsons (1980) for the generalization of Davidsonian decomposition to a verb's core arguments, but I believe the term "neo-Davidsonian", in description of such a system, is due to Dowty himself. Dowty's sketch of a neo-Davidsonian system is similar enough to Parsons' that I may add no further detail here. After exposition, Dowty reviews the criteria of *Completeness*, *Distinctness* and *Independence* within such a system.

As it turns out, under the neo-Davidsonian system, the properties of *Completeness* and *Distinctness* are built-in, because the only way to associate an argument with a verb is via a thematic-role predicate — semantically speaking, verbs no longer index their arguments at all. The bare verb merely describes an eventuality which may or may not be further constrained by adding the context of role players. Similarly, because each role is a separate predicate bearing its own truth conditions, the property of *Independence* is also satisfied without further stipulation in this system. However, if it is the case that the existence of an event entails the existence of a particular role player, then this information must be stipulated separately, via some sort of system of meaning postulates, such as the one in (70).¹⁴ That is, one difference between the two systems under examination is that the ordered argument system, by defining roles in terms of lexical entailments, already contains the fact that the existence of a *giving* event in the model entails the existence of an AGENT, because these entailments are part of the definition of AGENT. Contrastingly, in the neo-Davidsonian system, each verb must independently specify such entailments, if they are to be incorporated into the semantic system at all.

¹⁴For notes on modal logic symbolism see Garson 2018.

(70) $\forall e \square [\text{giving}(e) \rightarrow \exists x \text{ Agent}(x, e) \ \& \ \exists y \text{ Source}(y, e) \ \& \ \exists z \text{ Goal}(z, e)]$

(Dowty, 1989, 85)

What's more, because the system of lexical entailments which formed the basis for the thematic role type definitions under the ordered argument system is not yet constructed in this Neo-Davidsonian model, the system lacks collections of other entailments which were included in the ordered argument system as a basis, such as the fact that being the AGENT of *sings* entails the moving of vocal chords (but this is not entailed for the AGENT of *whistle*). These entailments which do not form a part of any of the L-thematic role type definitions of the ordered argument approach must be stipulated independently within the neo-Davidsonian system, as part of the lexical meanings of the verbs.

After comparing the two systems, Dowty concludes that the expressive power of the two models is equal, once these two types of meaning postulates are added to the neo-Davidsonian system. The empirical data which Dowty uses to contrast the two systems is the optionality of arguments and, somewhat less empirically, whether syntactically optional arguments imply the existence of semantic ones. If they do, the ordered argument approach provides a compact representation of this implication, by virtue of including the required semantic arguments within the predicate symbol itself. If they do not, then the neo-Davidsonian system, with its conjunction of independent predicates, naturally provides a semantic option to include a predicate when a syntactic reflection is attested, and not to say anything about it otherwise — i.e. it succinctly handles the problem of variable arity without losing the implication relations (which was one of Davidson's original points). Dowty considers this point in the context of English verbs alongside event-denoting nouns, that is, nominalized verbs such as the classic “destroy” ~ “destruction” examples. The specifics of the argumentation are too detailed to reproduce here, but from a very high-level, Dowty argues that the facts of English motivate a hybrid approach, drawing on the ordered-argument system for the core arguments of verbs, but using the neo-Davidsonian system for event-denoting nouns (i.e. nominalized verb constructions) and verbal adjuncts. Parsons, on the other hand, interprets the facts differently, pointing out that languages other than English

freely allow argument dropping even in verbal constructions and that even languages such as English can be used to express impossible situations. As an example of the latter, in dreams or works of fiction impossible situations such as “I was stabbed, but nobody stabbed me.” can be described by English, and the semantics of such utterances should be representable (Parsons, 1980, 98).¹⁵

I find Parsons’ arguments more impactful in the context of my own work. Firstly, in my Lushootseed grammar project, where I touch on semantics, it is focused on structural semantics, the syntactic part of semantics, the “who did what to whom” part of semantics. I haven’t been in the business of writing out lexical entailments which hold between lexical items.¹⁶ The hope, of course, is that the structural semantics captured by a syntactic grammar will be compatible with models and techniques used by lexical semanticists in their work, but these concerns are not my first order of business. So, as I understand it, the part of Dowty’s argument for maintaining an ordered argument approach to English verbs is relatively out-of-scope for me. On the other hand, Parsons’ point that languages differ in their constraints on what exactly is optional is directly relevant — Lushootseed allows argument dropping in third person contexts. Finally, the point about representing the semantics of impossible situations is an important one. The fact that a situation is impossible in the real-world is not one that should be encoded as a constraint on the expressive power of the semantics of natural language. More generally, it’s easy to confuse constraints on a model with constraints on a modelling language.¹⁷

Proto-roles Dowty’s work on the semantic foundations of thematic roles provides a definitional and foundational approach to the topic. The last section of Dowty 1989 contains some speculative

¹⁵Note as well that this observation may form a direct argument against the need for the kinds of meaning postulates which Dowty suggests must be made part of the lexical meanings of verbs.

¹⁶Emily Bender often uses a witticism (ultimately credited to Barbara Partee, (Bender, pc)) to emphasize the scope of semantic investigations in her course on syntax: “the meaning of life is *life*”.

¹⁷See the introductory remarks of Pollard and Sag 1994 for a concise but insightful account of how linguistic theory ought to relate to formal models. From another perspective, the pitfalls of confusing the qualities of the representation language with the qualities of the thing to be represented are distinctly illustrated in Eco’s (1993, Ch. 14) review of efforts by Enlightenment scholars Leibniz and Wilkins’ attempts to build “philosophical languages.”

discussion on why L-thematic roles exist (i.e. why they provide such a powerful descriptive generalization within and across languages). Dowty explores this from a language acquisition perspective, looking at how thematic roles may form the basis of a prototype theory which allows the bootstrapping of more concrete lexical meanings.

Dowty followed this chapter with a paper in *Language* (1991) refining his position, arguing that once it is accepted that thematic role types are not mutually exclusive in their constituent properties, one can boil the thematic role related phenomena down to two clusters of properties which he terms *Proto-Agent* and *Proto-Patient*. This hypothesis is explicitly connected to the prototype theory of Rosch and Mervis (1975) as well as the transitivity scale of Hopper and Thompson (1980).

The contributing properties for Proto-roles which Dowty provides (see Table 7.1) do align with the dimensions which make up Hopper and Thompson's transitivity scale (see Table 7.2), but provide a different viewpoint on them. The transitivity scale ranks the situations described by verbs with respect to their transitivity by, in part, talking about the role players. For example, high on the transitivity scale are the situations in which the A argument is high in potency, and those in which O is totally affected — correspondingly, volition and affectedness appear in Dowty's contributing properties under association with AGENT and PATIENT, respectively.

Agent	Patient
Volitional involvement in the event or state	Undergoes change of state
Sentience (and/or perception)	Incremental theme
Causing an event or change of state in another participant	Causally affected by another participant
Exists independently of the event named by the verb	Does not exist independently of the event, or not at all

Table 7.1: Properties which contribute to AGENT and PATIENT proto-roles (Dowty, 1991, 572)

Dowty uses these contributing properties as the basis for a universal linguistic hypothesis, which he formulates as a pair of principles connecting Proto-roles to argument selection criteria. I have reproduced Dowty's hypothesis (Ibid, p. 576) as (71). The hypothesis embodied in these

	HIGH	LOW
PARTICIPANTS	2 or more participants, A and O	1 participant
KINESIS	action	non-action
ASPECT	telic	atelic
PUNCTUALITY	punctual	non-punctual
VOLITIONALITY	volitional	non-volitional
AFFIRMATION	affirmative	negative
MODE	realis	irrealis
AGENCY	a high in potency	A low in potency
AFFECTEDNESS OF O	O totally affected	O not affected
INDIVIDUATION OF O	O highly individuated	O non-individuated

Table 7.2: The dimensions of Hopper and Thompson's (1980) transitivity scale

principles, especially the *Argument Selection Principle*, is one that can be returned to after examining the Lushootseed data in more detail.

(71) a. *Argument Selection Principle*: In predicates with grammatical subject and object, the argument for which the predicate entails the greatest number of Proto-Agent properties will be lexicalized as the subject of the predicate; the argument having the greatest number of Proto-Patient entailments will be lexicalized as the direct object.¹⁸

1. Corollary 1: If two arguments of a relation have (approximately) equal numbers of entailed Proto-Agent and Proto-Patient properties, then either or both may be lexicalized as the subject (and similarly for objects).
2. Corollary 2: With a three-place predicate, the nonsubject argument having the greater number of entailed Proto-Patient properties will be lexicalized as the direct object and the nonsubject argument having fewer entailed Proto-Patient properties will be lexicalized as

¹⁸Dowty (1991) notes a number of difficult cases for this Principle: lexical doublets (*buy* and *sell*), “nonstandard lexicalizations” (single verbs which flout the principle such as *receive*, *inherit*, *sustain*) and syntactically ergative languages (*inter alia*). For syntactically ergative languages, Dowty offers that the principle holds in its inverse and argues that such a fact provides a reason not to conflate the concept of proto-agent with that of a grammatical subject nor proto-patient with grammatical object. For nonstandard lexicalizations, Dowty suggests that they may simply be thought of as exceptions, revealing the non-categorical nature of the “principle”.

an oblique or prepositional object (and if two nonsubject arguments have approximately equal numbers of entailed P-Patient properties, either or both may be lexicalized as direct object).

- b. Nondiscreteness: Proto-roles, obviously, do not classify arguments exhaustively (some arguments have neither role) or uniquely (some arguments may share the same role) or discretely (some arguments could qualify partially but equally for both proto-roles).

Conclusion of Background on Event Based Semantics

Davidsonian event-based semantics, the abstraction over verbal events in such a way that the meaning of a verb is represented by multiple predicates providing constraints on a variable, forms the backbone of several approaches to semantic representation, including the common approaches used by HPSG authors (Pollard and Sag, 1994, 318). Parsons' system takes decomposition to the extreme, wherein verbal predicates themselves provide no argument slots for role players. Instead, role players are joined to verbs using thematic-roles as predicates. Dowty examines the details of thematic roles from a model-theoretic perspective, contrasting a system where thematic roles are primitives (as in Parsons) to one where thematic roles are epiphenomena observable as collections of lexical entailments which must hold of verbal arguments. In the next section, I examine the approach in several varieties of MRS, comparing the usage of event-based semantics in those varieties to Dowty's analysis of thematic role systems.

*7.2.2 Examining *MRS*

Having outlined the semantic literature from the introduction of event-based semantics to linguistic thematic role theory, I now turn to the reflection of these ideas which is found in the design of Minimal Recursion Semantics. Although the basics of MRS were introduced in an earlier chapter, the MRS description language in fact provides a family of options for representing sentential

semantics. MRS provides a structural guideline which may be realized in any number of ways in specific applications.¹⁹ Even within the DELPH-IN community, there are a number of MRS variants which are in use in particular grammars. This section reviews those options. I begin with MRS as deployed in the English Resource Grammar (Flickinger, 2000; Flickinger et al., 2014b,a), discussing it in the light of the preceding section on event-based semantics. I look at Robust MRS (Copestake, 2006, 2007a,b, 2009), an MRS variant which was proposed in order to bridge the output of shallow processing systems with deep semantic representations.

English Resource Semantics (ERS), Grammar Matrix MRS (GMRS)

The usage of MRS in the English Resource Grammar (ERG, Flickinger 2000) is called English Resource Semantics (Flickinger et al., 2014b, ERS) and is documented on a series of wiki pages available via the public Internet (Flickinger et al., 2014a). My project derives from the Grammar Matrix core-grammar and customization system. That project, in turn, derives in part from the ERG (Bender et al., 2002). Ongoing work within the DELPH-IN community generally aims to keep things compatible between the ERG and the Matrix core grammar and customization system, although there are often minor differences when a particular innovation in one project has not yet been ported into the other.

Because of its size and scope, the ERG is seen as an implemented grammar *par excellence*, and consequently, its usage of MRS is looked to as a standard from which one should deviate with caution. This advisory note is not based on an argument from authority, but because the number of interacting phenomena the ERG accommodates makes it a proving ground for would-be representations. What's more, ERG semantics have proven their utility for downstream applications.²⁰ Because my Lushootseed meta-grammar is forked from the Grammar Matrix codebase,

¹⁹For those familiar with it, an analogy between the XML specification (Bray et al., 2008) and the particular markup languages which are written using it (such as, *inter multi alia*, TEI (TEI Consortium, 2019), which describes textual artifacts, or BeerXML (Smith et al., 2003), which describes brewing recipes) may be useful in understanding the relationship of MRS to the variants I describe below.

²⁰Projects which have consumed ERS as input and topics thereof (list from Flickinger et al. 2016): Machine translation: Bond et al. 2011; Information extraction and QA: MacKinlay et al. 2009; Ontology extraction: Herbelot and Copestake 2006; Question generation: Yao et al. 2012; Entailment recognition: Lien and Kouylekov

something very similar to ERS is the *de facto* starting point for my own project. In order to conveniently label that “something very similar” and to provide an unambiguous discussion, I refer to the MRS as implemented in the Grammar Matrix core-grammar and customization system as GMRS, as it is Grammar Matrix source files which are under consultation in most of this section. Most of the time the concepts are similar or identical between the GMRS and ERS, but the types illustrated below are all from the Grammar Matrix core-grammar file, so technically I am illustrating from GMRS. An earlier version of GMRS is documented extensively in Flickinger et al. 2003.

ERS and GMRS are Davidsonian As acknowledged by Copestake (2009, 4) ERS and GMRS are fundamentally Davidsonian representations — in contrast to a neo-Davidsonian system (as described above). First, the elementary predications contributed by verbs provide an event variable which may be referred to by other predicates in modification, coordination, subordination or other grammatical relations.²¹ Secondly, the core arguments of each verb are specified within the EP as a series of features. The template for the main predicate of a transitive-verb as output by the customization system is shown in (72). Such a configuration is reminiscent of Dowty’s ordered argument system in which thematic roles, should they be considered, would have to be derived from the lexical entailments which can be examined across a lexicon of verbs and arguments. Indeed, Copestake (2009, §3), in reference to Dowty’s (1991) Proto-Agent/Proto-Patient hypothesis, points out that “determining argument labelling by the obliqueness hierarchy still allows generalisations to be made for all verbs.[...] If this is correct, then we can, for example, predict that the ARG1 of any predicate in a DELPH-IN grammar will not have fewer p-agt properties than the ARG 2 of that predicate. (p. 5,6)”

2014; Preprocessing for distributional semantics: Herbelot 2013; Detection scope of negation: Packard et al. 2014; Robot control interface: Packard 2014.

²¹Additionally, DELPH-IN grammars may provide event property constraints express grammatical notions such as tense, aspect, mood or sentential force. See §3.5.2.

$$(72) \left[\begin{array}{l} \textit{arg12-rel} \\ \text{PRED } _ \text{LEMMA_v_SENSE_rel} \\ \text{LBL } h \\ \text{ARG0 } \textit{event} \\ \text{ARG1 } \textit{ref-ind} \\ \text{ARG2 } \textit{ref-ind} \end{array} \right]$$

Another aspect of elementary predicates in ERS is that instead of attempting to harmonize the obliqueness hierarchy for all lexically related predicates, ERS labels particular “senses” of a predicate which have differing argument structures. So, for example, with respect to a verb like *boil*, which exhibits the causative-inchoative alternation, ERS provides distinct predicate symbols: `_boil_v_cause_rel(ARG0 event, ARG1 ref-ind, ARG2 ref-ind)` and `_boil_v_1_rel(ARG0 event, ARG1 ref-ind)` where the `THEME` of boiling is the `ARG2` of the former but the `ARG1` of the latter. There is, of course, an alternative approach available in which the causative and inchoative verb classes and their predicate structures are to be related by rule. For Copestake, this is an impractically difficult task to maintain in a grammar which is intended to scale: there are edge-cases everywhere and the utility of such a system is completely undermined if the labels are not consistent across all of the partially overlapping classes of a natural language lexicon (Copestake, 2009, 5).²²

In sum, the typical ERS representations found in Matrix-derived grammars are Davidsonian: predicates associate with full argument signatures for n -ary arguments (where n may be > 1). If a notion of semantic roles were to be required or imposed, they could be specified along the lines of the ordered argument approach in Dowty 1989. In ERS, predicate names are organized and classified according to senses, which allows a loose coupling between related senses, indicating the connection by virtue of a shared lemma without requiring a full derivational lexicon to map all variants from a shared base.

With respect to Lushootseed, this means that semantic contribution of the few underived

²²Levin (1993) provides a start on what a full inventory would look like (for English).

bivalent verbs would be represented as in (72). Indeed, this is the approach taken in my system: the verbal choices file contains a section for the underived bivalent intransitives (which are, nevertheless, indicated as transitive in the choices file — because the customization system’s notion of transitivity is merely a reflection of syntactic valence). This is presented in further detail in §7.3 below.

Intrinsic Variables Predicates in ERS are expected to follow the Characteristic Variable Property, a constraint on semantic representations introduced in (Oepen and Lønning, 2006)²³ in the context of MRS sembanking. Because computing the logical equivalence of variable assignment in an underspecified logical form is not straightforward (Oepen and Lønning, 2006, 1252), the authors propose a variable-free form of predicate-argument structure called Elementary Dependencies (EDS). These structures allow MRSes to be decomposed into “discriminants” which partition the set of analyses under consideration by an analyst constructing the treebank.

Copestake (2009, 3) refers to the Characteristic Variable Property as “one which emerged from working with large-scale constraint-based grammars” and relies on it as an assumption about the input of the mapping algorithm she outlines in Copestake 2009, §4.1 for converting Robust MRS (RMRS) structures to Dependency MRS (DMRS). These two MRS variants are discussed below.

The Characteristic Variable Property is defined as follows: for each elementary predication which is not representing a generalized quantifier, the feature `ARG0` bears a value which is not the `ARG0` of any other predication which is not representing a generalized quantifier. The exception for quantifiers is due to the fact that in ERS the quantifier relations by convention represent the variable they bind as the value of `ARG0`.²⁴

One way to think about the `ARG0` features of MRS EPs is in the context of semantic indices. Linguistic forms, in many frameworks, are associated with indices. Constraints on these indices

²³Referred to in this paper as a *distinguished variable*.

²⁴The Characteristic Variable Property is also referred to here: <http://moin.delph-in.net/DelphinTutorial/Formalisms>

are sometimes associated directly with syntactic rules, such as the distribution of anaphoric pronouns in the well-known Principles A, B and C of Chomsky's binding theory (Chomsky, 1981). One effect of the Characteristic Variable Property is that it ensures that each linguistic form "starts out" with a unique index. Thus, when part of the sentence meaning requires that two semantic indices corefer, the sentential semantics must declare them to do so. In sentences such as "the morning star is the evening star", the copula relation for this construction posits a variable equality between the `ARG0` of the first and second instances of the *star* predicate, ideally in a way which is interpretable by the algorithm which produces logical forms from MRS. Encoding Chomsky's binding principles similarly amounts to positing constraints on the unique, symbolic indices contributed by these `ARG0` features.

RMRS

Copestake 2006 presents a modified form of MRS intended to provide compatibility with shallow processing systems: Robust Minimal Recursion Semantics (RMRS). Copestake observes the tradeoffs between deep and shallow processing systems: deep, precision systems provide detailed information about an input which can be analyzed under a system of logical semantic representations, but their construction is expensive and they are usually not robust to the noise contained in natural text sources (a single typo may cause an entire parse failure); shallow, statistical systems, on the other hand, are robust. They provide some answer even given noisy input, but do not normally output the sort of information which can be used to compute a full logical representation.

RMRS aims to close the gap between the deep and shallow systems by modifying MRS for ease of use with partial information. This is very much in line with one of the key motivations which undergird the MRS family generally: using underspecification to facilitate a layered approach and the combination of partial information sources (Bos, 1995; Copestake et al., 2005). The key observation in the initial introductory paper for RMRS (Copestake, 2006) is that if the syntactic structure of an MRS is allowed to be underspecified, the output of a shallow analysis system such as a part-of-speech (POS) tagger can provide a partial MRS by filling out the skele-

ton of elementary predications which come from the lexical information that the POS tagger offers.

To this end, RMRS breaks out verbal dependencies into independent predicate conjuncts in a way which is reminiscent of the neo-Davidsonian system of Parsons. It does not, however, go so far as to add thematic role labels. Instead, when considered from the point of view of Dowty (1989), RMRS presents a bit of a hybrid between the Ordered Argument system and the Neo-Davidsonian one: verbal arguments are broken out into independent conjuncts (neo-Davidsonian) but RMRS rejects thematic role names as predicate symbols: indexing arguments numerically instead, as ARG_N. These ARG_N pseudo-predicates link an argument to a predicate via its label.

This is perhaps best illustrated with an example. In (73), I present a partial MRS-RMRS example pair from Copestake 2006 for the sentence “Every fat cat sat on some table”. (73a) provides the MRS²⁵ and (73b) provides the RMRS.

(73) a. RELS l0:every_q(x, h1, h2),
 l1:_fat_a_rel(x),
 l1:_cat_n_1_rel(x),
 l4:_sit_v_1_rel(e0 [TENSE: past], x),
 l4:_on_p_2_rel(e1, e0, y),
 l5:some_q(y, h6, h7),
 l6:_table_n_1_rel(y)
 HCONS qeq(h1, l1),
 qeq(h6, l6)

²⁵In most DELPH-IN grammars, the EP `fat_a_rel` would also have a unique eventuality variable as its ARG₀, providing compliance with the Characteristic Variable Property. This example was reproduced from Copestake 2006 without changing the argument structure of the predications.

b.	RELS	l0:_every_q(x0), l1:_fat_j(x1), l2:_cat_n(x2), l4:_sit_v_1(e3), l14:_on(e4), l5:_some(x5), l6:_table_n_1(x6)
	HCONS	qeq(h1, l1), qeq(h6, l6)
	IN-GS	in-g(l2, l1), in-g(l4, l14)
	ARGUMENT-RELATIONS ²⁶	RSTR(l0, h1), BODY(l0, h2), RSTR(l5, h6), BODY(l5, h7), ARG1(l4, x2), ARG1(l14, e3), ARG2(l14, x5),
	VARIABLE-IDS	x0 = x1, x1 = x2, x5 = x6

RMRS differs from ERS in the following ways:

1. Lexical EPs are all of arity 1: they introduce their characteristic variable and nothing more.
2. Unique labels: in normal ERS (and in GMRS) label-sharing indicates conjunction within the variable-scope tree. RMRS (version 1) uses the labels to associate arguments to lexical predications via pseudo predicates. Thus labels may no longer be shared among lexical

²⁶So-termed in Copestake 2007a.

predications to indicate conjunction within a variable binding scope. Instead, a list of pairings, called *in-g* (mnemonic for in-group), provides this information.

3. Explicit identification: because information may be added incrementally by distinct non-local sources (or even at differing layers of processing by different systems), variable names are always introduced with distinct indices and equalities which hold between them are asserted explicitly.

RMRS was modified in Copestake 2007a,b in that the notion of *anchors* was introduced and the *in-gs* were removed. This reorganization allowed for the return of the convention of ERS and GMRS wherein label-sharing indicates logical conjunction in the scope-tree. Then, the new *anchors* are used to tie together the pseudo-predicates. I have modified the example from Copestake 2006 to comply with the changes to RMRS of Copestake 2007a,b and present such in (74). The motivation for the removal of the binary *in-g* relations in favor of label sharing (and likewise, introducing the anchors for linking pseudo-predicates to events) is given in Copestake et al. 2005, §2, which puts forward a general argument for the avoidance of transitively equivalent binary relations in semantic representations: binary relations which transitively define a group produce spurious ambiguities in formal representations — a problem which is avoided by having each member of the group associate itself to an abstract tag or node-id.

```

(74)  RELS      l0:a0:_every_q(x0),
        l0:a0 RSTR(h1),
        l0:a0:BODY(h2),
        l1:a1:_fat_j(x1),
        l1:a2:_cat_n(x2),
        l4:a3:_sit_v_1(e3),
        l4:a3:ARG1(x2),
        l4:a4:_on(e4),
        l4:a4:ARG1(e3),
        l4:a4:ARG2(x5),
        l5:a5:_some(x5),
        l5:a5:RSTR(h6),
        l5:a5:BODY(h7),
        l6:a6:_table_n_1(x6)
      HCONS    qeq (h1, l1),
               qeq(h6, l6)
      VARIABLE-IDS  x0 = x1,
                    x1 = x2,
                    x5 = x6

```

Discussion RMRS, in some sense, provides a backed-off version of the neo-Davidsonian approach found in Parsons. It is neo-Davidsonian in its usage of unary lexical predicates which introduce variables alongside binary predicates that join them together. It is backed-off in that it does not require the usage of thematic role labels for the binary predicates, using numeric arguments instead (referencing an obliqueness hierarchy), the interpretation of which may be made verb-specific.

There are certain aspects of the RMRS system which are attractive for modeling Lushootseed

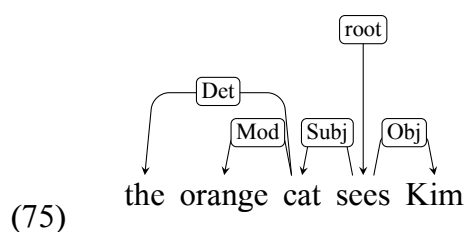
valence-increasing affixes. Under such a system, a valence-increasing affix might contribute a binary predicate which looks a lot like the *argument relations* of (73), conjoining a new variable with the existing lexical predicate contributed by the verb base. Such a system mirrors in the semantics the contribution pattern of the verb's morphology in a monotonically increasing way consistent with the rule-to-rule hypothesis of Bach 1976.²⁷ As a preview to a fuller discussion below, I note here that if the representation were allowed to abandon the Characteristic Variable Principle in such a way that the valence-increasing affixes share an ARG0 with the verb they attach to then the linking between the predicates is already accomplished without the use of additional anchors or label-sharing. However, because upholding the Characteristic Variable Principle allows the mapping of RMRS into DMRS (as well as EDS), I must review DMRS at least briefly in order to provide background for a discussion of an approach which loosens the Characteristic Variable Principle.

DMRS

The D in DMRS stands for Dependency. While the principles of Dependency Grammar (DG) stretch back, in some sense, to the earliest linguistic works, the contemporary theory is usually credited to Tesnière 1959 (Kruijff, 2002). From a high level, the idea of dependency grammar is to use directed arcs to directly indicate relationships with terminal elements. This is made more clear by contrasting dependency structures to those of constituency grammar. In a constituency grammar, particular grammatical elements may form the heads of phrases which may then be recursively combined. A phrase like *the orange cat*, under a constituency grammar, might be analyzed as an *NP* containing an internal \bar{N} , both headed by the noun *cat*. This same nesting of subphrases is also typically seen in verbal analysis: a verb may head a *VP* which is internal to an *S* which it also heads. In DG, these nestings are flattened out. Both *the* and *orange* are dependents of the head noun, but neither is internal to the other. Similarly, verb-headed

²⁷Questions about the application of the rule-to-rule hypothesis in the context of derivational morphology are important but beyond the scope of the present discussion. Also of note is that scholars differ in their answer to the question of whether or not these morphemes are derivational or inflectional (Beck, 2009, 542).

sentences with more than one argument phrase are analyzed according to directed arcs which point from the head to the dependent. No argument is more internal than another. Generally speaking, a DG analysis of a sentence will have the same number of nodes as there are words in the sentence (sometimes with a single additional abstract node representing the *ROOT* or *TOP*) whereas a constituency tree may contain arbitrarily more nodes than words, because of the intermediate structures described above.



From a scientific perspective, one aspect of DG is that it provides a hypothesis that the additional structures of constituency grammars are superfluous. Of course, in practice, whether something is superfluous depends on the task at hand. The hypothesis that certain structures may be eliminated is attractive to computational linguists who are already faced with a tractability problem when it comes to the implementation of practical tasks which rely on theoretical linguistic structures in their full, recursive glory. Dependency Grammars cut down the search space in parsing and generation. All else being equal, dependency-based NLP implementations are likely to be more efficient and performant than constituency based ones.

From this context²⁸ was born Dependency MRS (DMRS), an MRS variant based on dependency structures (Copestake, 2009). DMRS builds on the variable-free Elementary Dependencies Structures (EDS) of Oepen and Lønning (2006) in that it relies on the Characteristic Variable Principle to eliminate variables, but it extends the dependency graph to retain the semantic scope information which EDS do not. DMRS is described by translation from RMRS in the following

²⁸Copestake 2009, 6 cites “good engineering reasons” as a motivation for constructing a full-fledged dependency representation for MRS: she calls out specifically the “ease of readability” as well as “parser comparison and integration with distributional lexical semantics”. I have framed the context more broadly by introducing the principles of DG using syntactic examples, but I believe that the syntactic examples illustrate the simplification of structures (and the potential parsimony of analyses) which are hallmarks of dependency-based approaches in both semantics and syntax.

way. The starting point is the definition of three graphs which represent the structures of RMRS.

1. **label-eq-graph:** In RMRS each EP has a label, which may be shared. So the initial graph will contain a non-directional link between each pair of EPs which share a label.
2. **hole-label-graph:** RMRSs contain *qeqs* which map a hole to a label. Each label corresponds to n EPs, so there are n *qeq* links for a each hole.
3. **variable-graph:** Each variable is “introduced” (i.e. is the `ARG0`) of at most a single EP. So each non-hole argument which is filled by a variable is linked to the EP of which that variable is the `ARG0`. The pseudo-predicate `ARGN` relations of RMRS provide this information. In DMRS, the N becomes a label on a link between the EP whose argument position is filled and the EP which provided that variable as an `ARG0`.

These graphs are naively constructed as a starting point for the translation as the number of links can be reduced without loss of information. The variable graph provides a way to select the canonical links from the label-eq-graph so that redundant links can be eliminated. Instead of maintaining a link for each pair in the set, all elements of the set are linked via transitivity. The arc labels of the variable graph are modified to indicate their subsumption of a label-equality relationship. Copestake (2009, 6–8) outlines a further streamlining of these starting graphs into a compact representation which maintains both the scopal structures and the predicate argument structure.

DMRS is an important part of the DELPH-IN ecosystem. Since the introduction of DMRS, tools and procedures mapping other MRS flavors into DMRS have become available, including mapping from English Resource Semantics (Copestake et al., 2016). Perhaps more widely used than full DMRS structures are the EDS (which, like DMRS, provide a variable-free predicate-argument structure, but which do not provide scope information) which have been used in Se-

mEval²⁹ as well as CoNLL shared-tasks³⁰ Because of its proven usefulness in cross-framework semantic evaluation, maintaining compatibility with conversions into DMRS (or EDS) is an important consideration when crafting an MRS variant for a grammar.

As described above, one requirement for these conversions is that the grammar obeys the Characteristic Variable Property, the property of semantic representations wherein each semantic variable is the value of the `ARG0` feature of exactly one elementary predication. Informally, there is an intuition behind this property which corresponds to the notion that every bound variable referred to by the semantics should have a canonical source which is a linguistic element of the sentence.³¹ Formally, adherence to the CVP allows conversion to the variable-less format of DMRS to proceed by providing a decision criterion for selecting a semantic predicate which can stand-in for the variable in a variable-free representation. In the section on RMRS, I hinted at a representation which require a modification of the conversion algorithm in order to allow the `ARG0` variables of verbs to be linked to the additional predicates of valence-increasing morphology, I follow up on this hint in the next section, which presents an HPSG analysis of the valence increasing morphology in detail.

7.3 *Analysis*

In this section, I present an HPSG analysis of the Lushootseed phenomena described at the beginning of this chapter, taking into account the background literature reviewed. My approach draws on the ideas of Parsons as implemented in RMRS, but links pseudo-predicates to event-variables directly through `ARG0`-sharing. This approach avoids the additional *apparati* of the two

²⁹<http://sdp.delph-in.net/2015/>
<http://sdp.delph-in.net/2016/>

³⁰<http://mrp.nlpl.eu/>
<http://svn.nlpl.eu/mrp/2019/public/cfp.txt?p=27479>

³¹Bender (pc) points out that grammatical predicates, such as `compound_rel`, may introduce semantic variables and yet not be rooted in any particular lexical item. Such predicates form an exception to this informal generalization.

RMRS variants, eschewing both the addition of anchors and the repurposing of label-sharing (therefore, also avoiding the need for *IN-GS* to indicate variable scope). After stepping through the HPSG types and the modifications to the grammar generation system, I also discuss the modifications to the DMRS mapping algorithm which are required to support DMRS conversion for this analysis.

7.3.1 Overview

My analysis and implementation of valence-increasing morphology is motivated in part by the fact that, as Davis and Matthewson put it, “Salish verbs wear their derivational structure on their sleeves” (Davis and Matthewson, 2009, 1099). The transparently agglutinating structures which relate intransitives to transitives in Lushootseed are naturally attractive to the analytic grammarian. In accounting for these structures, I rely upon the principle of monotonic addition of information: rather than modifying existing predicates, my representations of valence-increasing morphs add new elementary predications the originals provided by the base. Taking RMRS as inspiration, I define a class of lexical rules which add pseudo-predications that provide a new argument and a new role to the eventuality indicated by the lexical verb. The new predication bears the name of the morphology which contributed it, allowing lexical semantics to happen in a subsequent stage of analysis, or not at all (depending on the application). I will take the first two sentences of LR1 Lesson Two as an illustrative example (76).

- (76) a. *ʔuʔuχ^w čəd*
 ʔu-ʔuχ^w čəd
 PFV-GO 1SG.SBJ
 ‘I went.’

- b. $\text{?u?u}\check{x}^w\text{tx}^w$ *čad* *ti* *čáčas*
 $\text{?u-?u}\check{x}^w\text{-tx}^w$ *čad* *ti* *čáčas*
 PFV-GO-ECS 1SG.SBJ DET child

‘I took the boy somewhere.’

Lushootseed Reader Volume 1, p.6

In both of the sentences of (76), the verb’s lexical entry contributes an elementary predication which looks like (77a): there is an ARG0, which is the verb’s internal index argument, representing the eventuality or situation described by the intransitive verb itself and there is an ARG1 slot, which represents the verb’s single role-player argument. After attachment of the perfective prefix, the value of the ARG0 feature will bear a constraint on the ASPECT feature. Then, in the intransitive example (76a), the syntactic rule which attaches the subject-marking auxiliary will provide additional constraints on the ARG1 feature value: setting [PER *1st*] and [NUM *sg*]. However, in the derivation of (76b), the morphological rule which attaches the $-\text{tx}^w$ suffix will add the elementary predication shown in (77b) to the semantics and co-index the ARG0 values between the two predications. This rule will also reorganize the argument structure linking so that the demoted subject, still semantically linked to the ARG1 of $\text{?u}\check{x}^w$, will be picked up as the first complement and the ARG1 value of the now introduced morphological predicate (76b) will be to linked to the index of the subject.

- (77) a. $\left[\begin{array}{l} \textit{arg1-relation} \\ \text{PRED} \textit{ “_?u}\check{x}^w_v_go_rel”} \\ \left[\begin{array}{l} \textit{event} \\ \text{TENSE} \textit{ tense} \\ \text{ASPECT} \textit{ aspect} \\ \text{MOOD} \textit{ mood} \end{array} \right] \\ \text{ARG0} \\ \text{ARG1} \textit{ ref-ind} \end{array} \right]$ b. $\left[\begin{array}{l} \textit{arg1-relation} \\ \text{PRED} \textit{ “_txw_x_ecs_rel”} \\ \text{ARG0} \textit{ event} \\ \text{ARG1} \textit{ ref-ind} \end{array} \right]$

The MRS in (78) shows a simplified view of the final constraints on the semantics of (76b). The intended interpretation of the two predications associated with the verb is akin to what was

seen in the discussion of RMRS, where the “pseudo” elementary predications collectively describe a semantic event, each contributing a role player. In this case, the totality of the linguistic information about the event is described by two sub-atomic predications: the first predication has a single argument-slot, which, in this example, indicates a goer, or better (perhaps), a $?u\check{x}^w$ -role; secondarily, this event has a type of causer, an ECS-role.

- (78) $\langle h_1, e_2 \{E.ASPECT \textit{PERFECTIVE}\},$
 $\left. \begin{array}{l} h_3: \textit{"?u\check{x}^w_v_go_rel"}(\textit{ARG0: } e_2, \textit{ARG1: } x_4 \{ \textit{PNG.PER 3RD} \}), \\ h_3: \textit{"txw_x_ecs_rel"}(\textit{ARG0: } e_2, \textit{ARG1: } x_5 \{ \textit{PNG.PER 1ST}, \textit{PNG.NUM SG} \}), \\ h_6: \textit{"ti_q_unique_rel"}(\textit{ARG0: } x_4, \textit{RSTR: } h_7, \textit{BODY: } h_8), \\ h_9: \textit{"\check{c}a\check{c}as_n_child_rel"}(\textit{ARG0: } x_4) \\ \{ h_7 =_q h_9 \} \end{array} \right| \rangle$

This approach is neo-Davidsonian, within a restricted domain of the grammar. It allows pseudo-predicates to provide additional role players to an event structure in a monotonically additive way, instead of hard-coding their combinatoric potential into the predicate symbols themselves. Following Parsons’ argumentation about the expressive abilities of language to indicate even situations which may be logically impossible, this approach does not hypothesize the existence of causers or other roles when they are not explicitly subcategorized for by the verbal morphology.

Since the vast majority of Lushootseed verbal radicals are monovalent, this approach may lead to the appearance that a neo-Davidsonian approach has been applied throughout. However, for the underlyingly bivalent verbs, my grammar represents their semantics directly as two place predicates, following standard ERS/GMRS practices. In this way, the overall grammar is actually Davidsonian, as ERS and GMRS are. In fact, the combinatoric potential of the neo-Davidsonian approach is constrained to verbal morphology — there are no syntactic rules affected. So all other areas of the grammar comply with ERS conventions.

Another item which merits note is the naming of the pseudo-predicates contributed by the

valence-increasing morphology. I chose to use the lexical content of the morphemes themselves to label the predicates. This is, in some sense, an approach which is congruent with the Ordered-Argument approach to indexing (which is standard in ERS). The predicate label itself provides the information which could be used to look up in an external resource which provides the lexical semantic entailments. The specific list of lexical entailments which can be stated of an external causer as opposed to an internal causer may, in fact, only be enumerable in the context of those morphemes once they have been paired with a primary predicate — this is how I envision the application of Dowty’s ordered argument system to my own neo-Davidsonian representation. In this way, my approach is a bit of a hybrid, but one that I believe provides all of the necessary information to consumers who may wish to carry out a lexical analysis of entailments within Lushootseed verbal structures. This agnostic approach to the lexical specifics is congruent with the idea that these morphemes represent derivational, as opposed to inflectional, morphology (Beck, 2009, 542) because one of the standard criteria for dividing between inflectional and derivational processes in languages is that the contribution of derivational morphemes is often context sensitive with respect to the base they attach to.

This overview section has provided the general idea of the approach. In the next section, I present the actual rules and type-hierarchy which I developed to implement this analysis.

7.3.2 Rules and Types

I begin by offering an apology for some of the names used in the analysis: the title of this case-study refers broadly to “valence-increasing morphology,” but the grammar customization subroutines and types described in this section refer instead to “transitivizers”. Not all of the valence-increasing affixes discussed here create transitives, however, in the morphological-marking sense of “transitive” which I offered in the first section of this chapter. The reason for this mismatch in terminology between this prose in this document and the type definitions presented as part of the analysis is due in part to the fact that the Grammar Matrix customization system itself uses the term “transitive” to indicate the property which I refer to as bivalence — that of taking two arguments. Verbs that take two arguments in the customization system source

code are referred to as transitive independent of the distinction in Lushootseed grammar³² which divides bivalent verbs into intransitive and transitive classes.

Marking a verb as “transitive” in the customization system allows the user to define constraints on both a subject and complement. In this way, using “transitivizing” in the rule names in the source code which implements my approach is consistent from the perspective of an engineer working on the codebase: the “transitivizing” lexical rules I define interact with Grammar Matrix core-grammar types which also bear the label “transitive”.

There are two aspects of the problem at hand which must be addressed: the organization of the lexical items which feed the rules and the organization of the rules themselves. The former is important because not all verbs are compatible with all of the affixes. The latter is important because the ontology of affixes under discussion shares certain properties at a high level and others at an intermediate level (causatives vs applicatives, etc).

I will first focus on the organization of the rules themselves, leaving allusions to a system of types which determines lexical compatibility, which is presented secondarily. Preliminary to both of these discussions is a presentation of the modifications to the Matrix core grammar’s notion of a transitive lexical item.

Valence Supertypes from the Core Grammar

Customization system choices files can mark verbal lexical types as `val=trans` or `val=intrans`. This specification indicates which valence related supertypes from the core-grammar to include on a particular type definition. The core grammar’s type for intransitive verbs works for my purposes out-of-the-box, it says (among other things) that the single argument of these items is headed by a noun. However, the corresponding type for transitives, which similarly declares the head-type of the first complement to be of type noun, is incorrect for Lushootseed since no Lushootseed sentence can bear more than a single direct NP argument. When the subject is omitted, or marked via an auxiliary, the object may be headed by a noun. However, for verbs

³²In fact, some authors have created clines of English bivalent verbs according to a semantic notion of transitivity. See Table 7.2 for the kinds of criteria that bear on the semanticist’s notion of transitivity.

such as the bivalent-intransitives, the object must be realized in a prepositional phrase. For this reason, I remove, redefine, then re-add, the type definition for *transitive-verb-lex* from the output grammar’s type definitions (79).³³

```
(79) def customize_lexicon(myLang):
    myLang.remove('transitive-verb-lex')
    typedef = """transitive-verb-lex := main-verb-lex & transitive-lex-item
    &
    [ SYNSEM.LOCAL.CAT.VAL.COMPS < #comps >,
      ARG-ST < [ LOCAL.CAT.HEAD noun ],
        #comps & [
          LOCAL.CAT [ VAL [ SPR < >,
            COMPS < > ],
            HEAD adp ] ] > ]."""
    myLang.add(typedef)
    gmcs/linglib/lushootseed/lexical_items.py
```

The definition for *transitive-verb-lex* in (79) will be assigned to the handful of Lushootseed verbs which are underlyingly bivalent. These verbs, such as *ʔətəd* “eat something”, *ʔulǎx*

³³It may be asked why the customization system’s default type for *transitive-verb-lex* is removed and then modified. That is, given that the metagrammatical system under discussion here is a fork which is customized to output Lushootseed, why I haven’t modified the code to insert what I want on the first pass, rather than removing and replacing in this way? The answer boils down to good engineering practice. I have attempted to add modifications to the existing customization system (and the Matrix core grammar) as little as possible, preferring instead to provide additional files and subroutines which are called after the existing ones. The advantage in this is twofold: by not interleaving my changes into the files of the project that I forked from, I can more easily merge in updates from the customization systems developers — as new functionality is incorporated in the upstream customization system, I can import that code and integrate it with mine much more easily than if all changes were interleaved. The second one is merely organizational: sprinkling changes to the customization logic throughout the existing codebase would quickly become unmaintainable — it would either cause existing regression tests to break or would require many ‘if/else’ branches to constraint the changes to Lushootseed customization, but finding all of these if/else constructions quickly becomes onerous. Instead, when I wish to make a modification to a subroutine in the existing system, I create a file of the same name in a Lushootseed namespace, and add my modifications there. This allows discoverability: all of the Lushootseed related changes are encapsulated in their own directory, and the naming convention indicates where in the customization system’s original layout the modifications are relevant.

Also of note, in the best case, existing customization system functionality is harnessed by my system, as opposed to duplicating this functionality with ad-hoc analyses and procedures. However, progress on a large project always requires a balance between ideal outcomes and intermediate goals, so the comment displayed reminds me to return to this type at a future date to look into the case-marking adpositions which are built into the case-marking library.

“gather something”, *pusil* “throw something”, subcategorize for two arguments: an agent as well as something which was acted upon. But crucially, for these verbs, the direct NP, when it appears, is interpreted as the subject — even when it is the only argument (see (80), which contrasts the bivalent intransitive with a verb which has undergone a transitivizing lexical rule). This may seem to be a violation of the rule of one-nominal interpretation. One rescue is to say that these are not true transitives, that they have no direct object slot available. In terms of practical implementation, we can simply say that the second argument of these verbs must be headed by an adposition. As shown in (79), the only noun-headed position available in the valence lists of these verbs is the subject’s.

- (80) a. *ʔuʔuləx̃axʷ čəd ʔə kʷi sʔuladxʷ tiʔəʔ*
ʔu=ʔuləx̃=axʷ čəd ʔə kʷi sʔuladxʷ tiʔəʔ
 IRR=gather=now 1SG.SBJ OBL DET salmon DEM

‘I will catch salmon there.’

Star Child (DS)

- b. *ʔugʷəčəd čəd tsi čáčas.*
ʔu-gʷəč-t čəd tsi čáčas
 PFV-search-ICS 1SG.SBJ DET.F child

‘I looked for the girl.’

LR1 p.10

The *transitive-verb-lex* type presented in (79) serves well for the bivalent intransitives which are lexically specified, but it will not work as written for intransitive verbs which have undergone a transitivizing lexical rule. For these, the system provides an alternative supertype, shown in (81). The two types *transitive-verb-lex* and *transitivized-verb-lex* differ in a few subtle ways. Firstly, *transitivized-verb-lex* inherits from *basic-two-arg-no-hcons* instead of *transitive-lex-item*. *transitive-lex-item* is a type from the Matrix core grammar (its definition is provided in (82), for reference) which links the indices of the two arguments into the single predicate found in LKEYS|KEYREL as ARG1 and ARG2. But, as discussed above, my analysis of these affixes is built upon monotonically building up a list of monovalent (pseudo-)predicates from the lexicon, not

modifying existing ones. So the ARG2 feature of *transitive-lex-item* is neither useful nor required in my analysis. Note as well that because *transitivized-verb-lex* is intended as a supertype for a family of rules, it does not fully constrain the head type of the complement.³⁴ The *+np* type is a disjunction type provided by the Matrix core grammar which is compatible with *noun* or *adp*. The reason for this underspecification is because in transitivized verbs, when the subject is realized via one of the subject marking auxiliaries, the object may be realized as a direct NP, not an oblique PP (80).

```
(81) transitivized-verb-lex := lex-rule & non-mod-lex-item & basic-two-arg-no-
      hcons &
      [ SYNSEM.LOCAL.CAT.VAL [ COMPS #comps & [ LOCAL.CAT.HEAD +np ],
                              SPR    < >,
                              SPEC   < >,
                              SUBJ   < #subj > ],
      ARG-ST < #subj . #comps > ].
```

```
(82) transitive-lex-item := basic-two-arg-no-hcons & basic-icons-lex-item &
      [ ARG-ST < [ LOCAL.CONT.HOOK [ INDEX ref-ind & #ind1,
                                    ICONS-KEY.IARG1 #clause ] ],
            [ LOCAL.CONT.HOOK [ INDEX ref-ind & #ind2,
                                    ICONS-KEY.IARG1 #clause ] ] >,
      SYNSEM [ LKEYS.KEYREL [ ARG1 #ind1,
                              ARG2 #ind2 ],
              LOCAL.CONT.HOOK.CLAUSE-KEY #clause ] ].
```

Having presented the lexical types which define both the underlying bivalent intransitives and transitivized verbs, I can turn now to the rule types.

³⁴The fact that *transitivized-verb-lex* is a supertype for a family of lexical rules breaks the Grammar Matrix naming convention in which items bearing the suffix *-verb-lex* should be lexical verbs, not lexical rule types. I think the break with convention is defensible in that these lexical rules are “sub atomic”, derivational rule types. What’s more, Bender (pc) points out that having lexical rule types inherit from these types, designed to construct lexical entries, may result in errors to do with composition. Further investigation of this point is left to future work.

Valence Increasing Rule Types

The output, or mother, of a lexical rule can differ almost completely from the inputs, or daughters. In HPSG implementations which offer defeasible constraints, all information is copied up unless otherwise indicated in the rule definition, but DELPH-IN TDL grammars do not offer defeasible constraints. Instead, rule authors must indicate which parts of the feature structure they wish to alter and what they wish to co-index between mother and daughter in a lexical rule. The Matrix core-grammar provides supertypes which offer convenient names for these coindexations for many classes of lexical rules. These convenience supertypes are found in the leaves of a hierarchy built from supertypes that “copy up” individual feature values. Common collections of these features are bundled into the convenience types such as *same-local-lex-rule*, which is shown alongside the supertypes which comprise it in (83).

```
(83) same-ctxt-lex-rule := lex-rule &
      [ SYNSEM.LOCAL.CTXT #ctxt,
        DTR.SYNSEM.LOCAL.CTXT #ctxt ].

      same-cont-lex-rule := lex-rule &
      [ C-CONT [ HOOK #hook ],
        DTR.SYNSEM.LOCAL.CONT [ HOOK #hook ]].

      same-agr-lex-rule := lex-rule &
      [ SYNSEM.LOCAL.AGR #agr,
        DTR.SYNSEM.LOCAL.AGR #agr ].

      same-cat-lex-rule := lex-rule &
      [ SYNSEM.LOCAL.CAT #cat,
        DTR.SYNSEM.LOCAL.CAT #cat ].

      same-local-lex-rule := same-cat-lex-rule &
                           same-cont-lex-rule &
                           same-ctxt-lex-rule &
                           same-agr-lex-rule.
```

Transitivizing lexical rules, in my analysis, will require changes to valence lists and will provide new semantic information in *C-CONT*. The nearest out-of-the-box supertype provided by

the Matrix core-grammar for such a structure is the *val-change-only-lex-rule*. This rule type inherits (indirectly) from *no-ccont-lex-rule*, however, and therefore will not suffice. For that reason, my system provides a new type which allows both changes to valence lists as well as a non-empty *C-CONT* feature, coindexing all of the other structures between mother and daughter (*val-change-with-ccont-lex-rule* in (84)).

With *val-change-with-ccont-lex-rule* defined, a type-definition for a supertype (which will subsume all of the transitivizing lexical rules) can be added to the grammar by conjoining *val-change-no-ccont-lex-rule* with *transitivized-verb-lex* (discussed above) and *infl-lex-rule* (which indicates that the lexical rule bears phonological content). Finally, a constraint on the *DTR* feature ensures its position class in the morphological pipeline (84).

```
(84)  val-change-with-ccont-lex-rule := same-light-lex-rule &
                                           same-modified-lex-rule &
                                           same-non-local-lex-rule &
                                           same-head-lex-rule &
                                           same-hc-light-lex-rule &
                                           same-posthead-lex-rule &
                                           same-mc-lex-rule &
                                           same-ctxt-lex-rule &
                                           same-agr-lex-rule &
                                           no-icons-lexrule.

transitivizing-lex-rule-super := infl-lex-rule &
                                val-change-with-ccont-lex-rule &
                                transitivized-verb-lex &
                                [ DTR aspect-lex-rule-super ].
```

One thing to note about the type *transitivizing-lex-rule-super* is that the type name itself provides a moment of coordination with the position classes specified by the choices file. Specifically, when the morphotactics library processes a position class named *x*, a lexical rule type is created in the output grammar named *x-lex-rule-super* which is supertype to all of the lexical rule types defined under that position class. Because the Lushootseed transitivizing morphology is (for the most part³⁵) in complementary distribution, my morphotactic choices collects the

³⁵The exceptions are the secondary suffixes. However, as noted above, since these secondary suffixes nearly

lexical rule types under a single position class called *transitivizing*. Thus, the built-in morphotactics system generates a type name *transitivizing-lex-rule-super* which provides a supertype for all lexical rule types which fall into this position class. I can then provide my additional functionality for the type *transitivizing-lex-rule-super* as shown in (84) and rely on the customization system's TDL processing to merge the additional constraints with those provided by the morphotactics system before the type is serialized for inclusion in the output files.

With this rule supertype in place, I can now add more specific types which implement a still relatively generic version of causative and applicative rules. A template is created as a Python formattable string, which allows me to accommodate either ERS or Lushootseed MRS predicate format, based on the choices file. The type generation logic (and template) is shown in (85).³⁶

always require the presence of ics marking, I treat the entire complex as just another one of these transitivizers. There are actually a few other idiosyncratic aspects of secondary suffixes which must be considered as well. These are presented below.

³⁶I have shown the Python grammar generation code here and in a number of subsequent examples in order to present the logic of grammar generation alongside the types generated themselves. In several contexts, I take advantage of Python's string formatting syntax to parameterize the type which will be output in a succinct way. Specifically, the syntax {} in a string object such as `base_typedef` in (85) indicates where text will be inserted when `base_typedef.format(predargs)` is called. The full syntax for this formatting language can be found online at <https://docs.python.org/2/library/stdtypes.html#str.format>.

```
(85) def add_causative_supertype(mylang, choices):
    comment = """;;; add a causer, indexed with new arg on subj
    ;;; need to re-enforce fully saturated subject""
    mylang.add_literal(comment)

    base_typedef = ""causative-lex-rule-super := lex-rule &
    [ C-CONT [ RELS <! [ {} ] !>,
      HOOK [ INDEX #e,
            XARG #x ] ],
    ARG-ST < [ LOCAL [ CAT [ HEAD noun,
                          VAL [ SPR < >,
                                COMPS< > ] ],
              CONT.HOOK.INDEX #x ] ],
      #oldsubj >,
    DTR [ SYNSEM.LOCAL [ CAT.HEAD verb,
                        CONT.HOOK [ INDEX #e,
                                    LTOP #h,
                                    XARG #intern-subj ] ],
      ARG-ST < #oldsubj & [
        LOCAL.CONT.HOOK.INDEX #intern-subj ] > ] ].""

    if choices.get('semantics') == 'emrs':
        predargs = 'LBL #h, ARG0 event, ARG1 event & #e, ARG2 #x'
    else:
        predargs = 'LBL #h, ARG0 event & #e, ARG1 #x'
    typedef = base_typedef.format(predargs)
    mylang.add(typedef)
```

The causative supertype shown in (85) does a few things of note: first, it adds a new predicate to the RELS list. This new predicate either shares an ARG0 with the main lexical predicate, or it takes the ARG0 of the main lexical predicate as its ARG1 (as noted above). The ARG1 is linked to the XARG of the C-CONT|HOOK (which will be indexed with SYNSEM|LOCAL|HOOK by virtue of an ancestor supertype *norm-sem-lex-item*). Secondly, it pushes a new dependency onto the first position of the mother’s argument structure list. The old subject is copied up wholesale, but is now the first complement. Thirdly, it copies up the DTR|SYNSEM|LOCAL|HOOK|INDEX so that the main lexical predicate still provides the index.

This supertype is shared by the full causatives (ICS and ECS) as well as the “semi-causatives” (CSMD, ACT) which form derived bivalent intransitives. The difference between these two sub-

types is the argument realization patterns: the bivalent intransitives ensure that the second argument may only be realized as an oblique PP, while the true causatives mark their subject as OPT +, so that any NP which is picked up must be interpreted as a complement (86). With these supertypes in place, I can add a series of types to provide the actual predicate labels as leaf types in this hierarchical analysis (87).

```
(86) causative-lex-rule := causative-lex-rule-super &
      [ ARG-ST.FIRST.OPT + ].
```

```
      bival-intr-causative-lex-rule := causative-lex-rule-super &
      [ ARG-ST < [ ], [ LOCAL.CAT.HEAD adp ] > ].
```

```
(87) altv-lex-rule := applicative-lex-rule &
      [ C-CONT.RELS <! [ PRED "_c_x_altv_rel" ] !>
      SYNSEM.LOCAL.CAT.HEAD altv-compat ].
```

```
      csmd-lex-rule := bival-intr-causative-lex-rule &
      [ C-CONT.RELS <! [ PRED "_b_x_csmd_rel" ] !>,
      DTR.SYNSEM.LOCAL.CAT.HEAD csmd-compat ].
```

```
      act-lex-rule := bival-intr-causative-lex-rule &
      [ C-CONT.RELS <! [ PRED "_alikw_x_act_rel" ] !>,
      DTR.SYNSEM.LOCAL.CAT.HEAD act-compat ].
```

```
      ics-lex-rule := causative-lex-rule &
      [ C-CONT.RELS <! [ PRED "_t_x_ics_rel" ] !>,
      DTR.SYNSEM.LOCAL.CAT.HEAD ics-compat ].
```

```
      ecs-lex-rule := causative-lex-rule &
      [ C-CONT.RELS <! [ PRED "_txw_x_ecs_rel" ] !>,
      DTR.SYNSEM.LOCAL.CAT.HEAD ecs-compat ].
      mylang.add(typedef)
```

```
      dc-lex-rule := causative-lex-rule &
      [ C-CONT.RELS <! [ PRED "_dxw_x_dc_rel" ] !>,
      DTR.SYNSEM.LOCAL.CAT.HEAD dc-compat ].
```

The applicative rule types are handled similarly: there is a template string which can be formatted according to the semantic arguments desired, and this is then inherited by the rule types which add the appropriate predicate labels (88).

```
(88) base_typedef = ""applicative-lex-rule := same-cont-lex-rule &
      same-lkeys-lex-rule &
      [ C-CONT.RELS <! [ {} ] !>,
        SYNSEM.LOCAL.CONT.HOOK.INDEX #e,
        ARG-ST < #subj & [ OPT + ], [ LOCAL [ CAT.HEAD +np,
                                                    CONT.HOOK.INDEX #x ] ] >,
        DTR.ARG-ST.FIRST #subj ].""

      if choices.get('semantics') == 'emrs':
        predargs = 'ARG0 event, ARG1 event & #e, ARG2 #x'
      else:
        predargs = 'ARG0 event & #e, ARG1 #x'
      typedef = base_typedef.format(predargs)
      mylang.add(typedef)

      typedef = ""altv-lex-rule := applicative-lex-rule &
      [ C-CONT.RELS <! [ PRED "_c_x_altv_rel" ] !>
        SYNSEM.LOCAL.CAT.HEAD altv-compatible ].""
      mylang.add(typedef)
```

In (88), the new argument is appended to the argument structure list, rather than pushed onto the beginning, as expected for applicatives. This rule differs from the causative supertype in two other ways, however. First, the applicative supertype is able to identify the `CONT` feature of the mother and daughter via *same-cont-lex-rule*. Because no changes to the `XARG` or other `HOOK` features are required and the new information comes in via `C-CONT`, there no need to reformulate the `CONT` value on the mother beyond the append of new information. Also of note is the supertype *same-lkeys-lex-rule*. This supertype creates an identity constraint between `SYNSEM|LKEYS` and `DTR|SYNSEM|LKEYS`.

In fact, my meta-grammar system adds this constraint to all of the lexical rule types except the causatives (89). The `LKEYS` feature provides a convenience feature for accessing semantic information provided by the lexical entry, so that lexical rules which provide constraints on this

information do not have to manage the length of the `RELS` list directly. The `LKEYS` feature relates to the transitivizing lexical rules in two ways: First, the transitivizing lexical rules may apply after other lexical rules,³⁷ and they require access (although this access is indirect, through the `DTR|HOOK` features) to the `LKEYS` feature as provided by the original lexical entry (in order to properly link the predicate they provide to the lexical one). Second, the transitivizing lexical rules cannot themselves enforce the identification of `LKEYS`, because they change the linking between embedded `XARG` and the new `XARG`, and if the entire structure were identified, this would effectively create a constraint which would implement a reflexive (the new argument would be identified with the existing one).

The applicative lexical rules, on the other hand, make no changes to the linking into the `KEYREL`. Because they add an argument without changing the linking of an existing one, they can inherit from *same-lkeys-lex-rule*.

```
(89) def copy_lkeys(mylang, choices):
    # Copying up of lkeys is required for the transitivizing lexical rules
    # to
    # apply to the output of another rule.
    typedef = """same-lkeys-lex-rule := lex-rule &
    [ SYNSEM.LKEYS #lkeys,
      DTR.SYNSEM.LKEYS #lkeys ]."""
    mylang.add(typedef)

    for vpc in choices.get('verb-pc'):
        rule_super = vpc.get('name')
        if rule_super and rule_super != 'transitivizing':
            typedef = rule_super + '-lex-rule-super := same-lkeys-lex-rule.'
            mylang.add(typedef)
```

The preceding paragraphs have detailed the approach to the rule system of my analysis for both the applicatives and causatives. The Lushootseed data in the testsuite also contains exam-

³⁷My metagrammar attaches prefixes before suffixes. This choice was arbitrary — I could have engineered the position classes to be processed in another order. In future work, I will likely reorganize both the FST structures and the morphotactic processing in the syntactic grammar to something more along the lines of what Czaykowska-Higgins (1996) has for Nxạmxcín [col] (as appropriate for Lushootseed).

ples including one of the secondary affixes, the dative marker or *-yi-t* forms. These datives are trivalent, in the grammar source code they are treated as a class of ditransitives. The approach is similar to, yet distinct from the approach the simple transitives. I begin by providing a supertype to constrain the output's argument realization pattern, shown in (90).

```
(90) di-transitivized-verb-lex := lex-rule & basic-three-arg-no-hcons &
      [ SYNSEM.LOCAL.CAT.VAL [ COMPS #comps,
                              SPR < >,
                              SPEC < >,
                              SUBJ < #subj > ],
      ARG-ST < #subj . #comps > ].
```

Equivalently to the approach for the generic supertypes described above, I use a Python formattable string to define the supertype in a way that can be conveniently output in either ERS style or Lushootseed MRS style. This is similar to the transitives in certain ways, but the notable difference is that these *-yi-t* rules must be provided in two flavors, one which applies to transitives and one which applies to intransitives (91).

```

(91) def add_yid(mylang, choices):
    yid_predargs = 'ARG0 #e & event, ARG1 #x1'
    ics_predargs = 'ARG0 #e & event, ARG1 #x2'
    if choices.get('semantics') == 'emrs':
        yid_predargs = 'ARG0 event, ARG1 #e & event, ARG2 #x1'
        ics_predargs = 'ARG0 event, ARG1 #e & event, ARG2 #x2'

    typedef = ""yid-super := di-transitivized-verb-lex & val-change-with-
ccont-lex-rule &
[ C-CONT [ RELS.LIST.FIRST [ PRED "yid_x_dat_rel",
                            {} ] ],
  SYNSEM [ LOCAL.CONT.HOOK.INDEX #e,
            LKEYS.KEYREL [ ARG0 #e,
                           ARG1 #oldxarg ] ],
  ARG-ST < #oldsubj &
    [ OPT + ],
    [ LOCAL [ CAT [ HEAD noun,
                  VAL [ SPR < >,
                       COMPS < > ] ],
              CONT.HOOK.INDEX #x1 ] ],
    [ LOCAL [ CAT.HEAD adp,
              CONT.HOOK.INDEX #oldxarg ] ] >,
  DTR.ARG-ST.FIRST #oldsubj ]."".format(yid_predargs)
    mylang.add(typedef)

    typedef = ""yid-itr-lex-rule := yid-super &
[ C-CONT.RELS [ LIST.REST [ FIRST [ PRED "ics_x_rel",
                                    {} ] ],
                REST #last ],
  LAST #last ],
  SYNSEM.LOCAL.CONT.HOOK.INDEX #e,
  ARG-ST.FIRST.LOCAL.CONT.HOOK.INDEX #x2 ]."".format(ics_predargs)
    mylang.add(typedef)

    typedef = ""yid-tr-lex-rule := yid-super &
[ C-CONT.RELS [ LIST.REST #last,
                LAST #last ],
  ARG-ST < [ ], [ ], #oldcomp >,
  DTR [ ARG-ST < [ ], #oldcomp > ] ].""
    mylang.add(typedef)

```

Some of the data which motivates these two variants is shown in (92). In the first example, *-yi-t* combines with the intransitive *ʔuχʷ*, but in the second example, the form which feeds *-yi-t*

is already transitivized, via the external causative. Furthermore, (93) shows that the output of *-yi-t* is trivalent, even when it is applied to an intransitive.

(92) a. *ʔuʔuχ^wyi-d* *ti* *čáčas*.

ʔu-ʔuχ^w-yi-t *ti* *čáčas*

PFV-GO-DAT-ICS DET child

‘[Someone] went instead of the boy.’

b. *ʔuʔuχ^wtx^wyi-d* *ti* *čáčas*.

ʔu-ʔuχ^w-tx^w-yi-t *ti* *čáčas*

PFV-GO-ECS-DAT-ICS DET child

‘[Someone] took [something/someone] somewhere for the boy.’

LR1 p.34

(93) *ʔuʔabyid* *čəd* *ti* *čáčas* *ʔə* *ti* *sq^wəbay?*

ʔu-ʔab-yi-t *čəd* *ti* *čáčas* *ʔə* *ti* *sq^wəbay?*

PFV-extend-DAT-ICS 1SG.SBJ DET child OBL DET dog

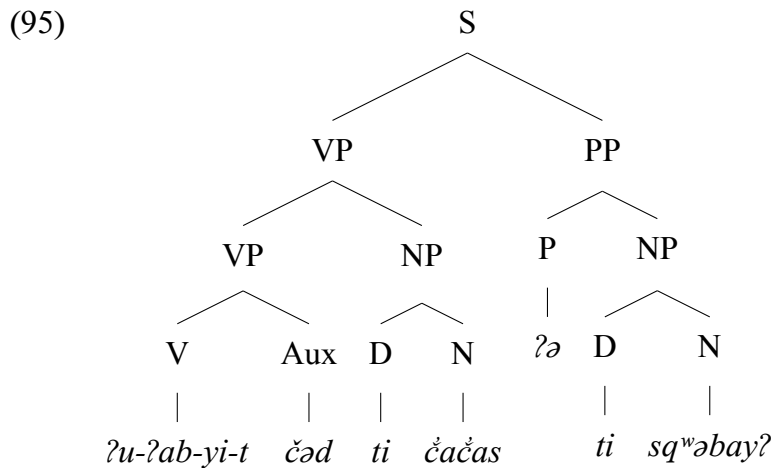
‘I gave the dog to the boy.’

LR1 p. 36

I accommodate the two inputs as shown in (91), by providing two forms of *-yi-t*, an intransitive, which posits two new valence dependencies and two new relations, an ics relation, reflected by the *-t*, and the dative relation. The new causative argument is added as the subject and the dative argument is the first complement — in that way, the *-yi-t* rule is applicative-like. The existing argument is demoted to an obliquely marked second complement. Example (94) shows a the current system’s Lushootseed MRS output when analyzing the sentence in (93): there are three verbal relations associated with *ʔuʔabyid*, an extended thing, which is *ti sq^wəbay?* (“the dog”), a causer, which marked as 1sg and a dative argument, which is *ti čáčas*, the boy. For context, a simplified view³⁸ of the syntactic derivation tree is shown in (95).

³⁸The tree-structure shown here is a simplification in that descriptive abbreviations are used for node labels. In

- (94) $\langle h_1, e_2 \{E.ASPECT \textit{PERFECTIVE}\},$
 $h_3: \textit{"_?ab_extend_rel"}(\textit{ARG0}: e_2, \textit{ARG1}: x_4 \{PNG.PER \textit{3RD}\}),$
 $h_3: \textit{"_yid_x_dat_rel"}(\textit{ARG0}: e_2, \textit{ARG1}: x_5 \{PNG.PER \textit{3RD}\}),$
 $h_3: \textit{"_ics_x_rel"}(\textit{ARG0}: e_2, \textit{ARG1}: x_6 \{PNG.PER \textit{1ST}, PNG.NUM \textit{SG}\}),$
 $h_7: \textit{"_ti_q_unique_rel"}(\textit{ARG0}: x_5, \textit{RSTR}: h_8, \textit{BODY}: h_9),$
 $h_{10}: \textit{"_čacás_n_child_rel"}(\textit{ARG0}: x_5),$
 $h_{11}: \textit{"_ti_q_unique_rel"}(\textit{ARG0}: x_4, \textit{RSTR}: h_{12}, \textit{BODY}: h_{13}),$
 $h_{14}: \textit{"_sqʷəbay?_n_dog_rel"}(\textit{ARG0}: x_4)$
 $\{ h_8 =_q h_{10}, h_{12} =_q h_{14} \} \rangle$



In the case of the transitivized input, the transitive-taking variant of the *-yi-t* suffix only assigns a new dative role and dependency — the existing causative relation and argument dependency is kept as created by the transitivizing lexical rule.

Verbal Lexicon

Because Hess (1973) alludes to a grammaticalized distinction between patient-oriented and agent-oriented (cf “experiencer oriented”, above) verb radicals, my original implementation of the verbal lexicon divided verb roots into two head types correspondingly. The distinction

the actual representations provided by the grammar, the nodes of the tree are full feature structures.

which is implied by Hess (1973) is that patient oriented verbs can take the internal causative, while agent oriented verbs take the external causative and allative applicative. So in my original implementation, I encoded these verbal head types as constraints on the internal causative and allative applicative rules, reflecting my understanding of Hess (1973).

When implementing further causative types, such as the diminished control causative, I saw that the pattern I had established was problematic: the diminished control causative took verbs from both the agent-oriented and patient-oriented classes. I saw a similar thing when I implemented the middle-voice causatives: both of these two new affixes cross cut my original classes. As an interim measure, I placed only the generic *verb* constraint on the input of these rules. This worked to capture the existing data, but didn't protect against overgeneration pairing verbs with these affixes where the combination would be ungrammatical. I also began to run across particular verbs in the patient-oriented class which could take the allative applicative. Via personal communication, Zahir indicated that he knew of no reason in principle why such a combination would be ungrammatical in general. So I began to reexamine the notion of patient versus agent orientation as the basis for verbal morphology classes.

The nail in the coffin for my original semantic-orientation-based approach to verb classes came when I ran across *hədiw* ("go inside"). As Hess notes in LR1 (p. 16), "Usually a verb taking *-tx*" does not also take *-d*. ...However, *?uhədiw* is such a verb." An inevitable moment in grammar development: I had hit the point where I needed a system for lexical exceptions to the general pattern, and this prompted a refactoring of my analysis along new lines.

In lieu of a proper characterization of the exact lexical classes which Lushootseed verbs fall into,³⁹ I opted for an approach which would maximize flexibility. At the worst case, where no generalization is possible, the number of verb classes may be equal to 2^n where n is the number of affixes which a verb may be compatible or incompatible with. My starter grammar implements six transitivity lexical rules: ICS, ECS, DC, ACT, CSMD, ALTV. While 64 classes is not an unmanageable number, it is already more classes than I wish to manage "by hand", and

³⁹Hess (1993) suggests the outlines of such a system, but the inventory listed is only a handful of verbs.

of course, given this exponential function, the number of classes would become quite unwieldy should I end up implementing just one or two more affixes.

My approach, then, is to create a feature structure which manages morphological compatibility as an array of flag features as shown in (96).⁴⁰ This feature structure is appropriate for the *head* feature-structures of type *verb*, and allows lexical entries to say which rules they are compatible with. The metagrammar generation system also creates convenience types which allow the descriptive setting and unsetting of these flag features (97). For common bundles of flag settings, I also create conjunction types which set a particular collection of flags to + and the rest to – (97). With these types in the grammar, the morphological rule instances only constrain their application to the relevant *-compat* type, as seen in (87).

```
(96) morph-compat := avm &
      [ ICS luk,
        ECS luk,
        DC luk,
        ACT luk,
        CSMD luk,
        ALTV luk ].
```

```
verb +: [ MORPH-COMPT morph-compat ].
```

```
(97) ics-compat := verb & [ MORPH-COMPAT.ICS + ].
      ecs-compat := verb & [ MORPH-COMPAT.ECS + ].
      ...
      ics-incompat := verb & [ MORPH-COMPAT.ICS - ].
      ecs-incompat := verb & [ MORPH-COMPAT.ECS - ].
      ...
      altv-ecs-only := altv-compat & ecs-compat &
                      ics-incompat & dc-incompat & act-incompat & csmd-incompat.
```

The remaining matter, then, is to organize the verbs in the choices file to use these types. As shown in (98), this merely requires using the relevant convenience type name as the verb's

⁴⁰This approach is similar to the work of the ALTS feature of the ERG (Flickinger, 2000) as well as the flag features of the morphotactics system of Goodman (2013).

type name in the choices file: the type-name for verb5 coordinates these morphological constraints with these verbs. Then, in order to accommodate one-offs such as *hədiw* without having to create new classes in the choices file and new collection types in the grammar generation system, I added the functionality to allow specific lexical entries within a type to override the constraints of their type. I did this using a customization system feature `mcadd` (for “morphological-compatibility addition”). When `mcadd` is specified on a lexical entry in the verbal choices file, the customization system computes the *morph-compat* type which will be required by adding the lexical rules specified in the value to those of the class. In the case of *hədiw*, the type would be *altv-csmd-ecs-ics-only*. If the grammar does not yet have such a type, it is created (by joining the relevant convenience types) and then the particular lexical entry is created under this new class instead of the one the choices file specifies directly. The complementary `mcdel` could be implemented as well, but I have not yet done so in my grammar generation system. The advantage of this system for managing morphological compatibility is that it allows me to organize the lexical choices file in convenient ways: verb1 and verb2 are still associated with agent-oriented and patient-oriented semantics (respectively), and I can still capture the broad generalization such as the fact that *most* verbs I have run across which are compatible with the allative applicative are also compatible with the external causative. But under this approach, the organizational structure of the choices file doesn’t require me to manage an intractable number of verb classes as I run across particular verbs which are attested in idiosyncratic constructions.

```

(98)  verb5_name=altv-ecs-only
      verb5_supertypes=verb1
      verb5_stem1_orth=?uǰw
      verb5_stem1_pred=?uǰw_v_go_rel
      verb5_stem2_orth=?ǰǰ
      verb5_stem2_pred=?ǰǰ_v_come_rel
      verb5_stem3_orth=?tuǰw
      verb5_stem3_pred=?tuǰw_v_go.home_rel
      verb5_stem4_orth=?talil
      verb5_stem4_pred=?talil_v_go.ashore_rel
      ...
      verb5_stem22_orth=?šut
      verb5_stem22_pred=?šut_v_see_rel
      verb5_stem22_feat1_name=mcadd
      verb5_stem22_feat1_value=dc
      verb5_stem23_orth=?həd?iǰw
      verb5_stem23_pred=?həd?iǰw_v_inside_rel
      verb5_stem23_feat1_name=mcadd
      verb5_stem23_feat1_value=csmd,ics

```

This section has presented my current system of lexicon management. It is a system of convenience which allows flexibility in the granularity of application of particular observations. I can create general rules and exceptions to those rules with relatively simple changes to the choices file. This also concludes my presentation of my current analysis and the implementation thereof. In the next section, I turn to the ramifications of my analysis on the mapping from full MRS into dependencies.

7.3.3 *DMRS mapping*

In initial presentations of this analysis, several researchers within the DELPH-IN community expressed concerns about the incompatibility of the approach to the DMRS mapping algorithm of Copestake (2009). In this section, I offer a reply to these concerns in three ways. Firstly, I present an alternative analysis which does satisfy the Characteristic Variable Principle and show how a trivial mapping can be implemented to swap back and forth between the representation I presented above, and this alternative. Secondly, I defend my own analysis as preferable by virtue

of containing the stronger linguistic hypothesis (as well as in reference to certain arguments about linguistic productivity). Thirdly, I show how a simple modification of the DMRS algorithm Copestake (2009) can accommodate my original analysis.

An alternative: more variables

The alternative analysis posits more event-type variables. Instead of contributing a predicate with two argument positions (ARG0 and ARG1), they can introduce a predicate with three positions. Instead of sharing their ARG0 with the lexical verb base they attach to, they can take this value as ARG1, contributing their own, unique ARG0.

This analysis was added to the grammar generation system as a choice of semantics. When the input choices file specifies `semantics=emrs` the output grammar is created with the additional variable slot in the valence-increasing affixes. The implementation detail of this choice is included above in the listings in: (85), (91) and (88).

A stronger hypothesis

I argue that the evidence at hand underdetermines the solution to the principled question of how many event type variables ought to be present in the semantics of these constructions. Given the lack of evidence, it is preferable to constrain the semantic output, and extend it when evidence comes available which requires the extension.

Let me again return to the examples in (76), which contrasts the verb base form of the verb $?u\check{x}^w$ (“to go”) with the transitivized $?ux^wtx^w$ (“bring someone”). Under one analysis, there is only one event variable and the external causative marker $-tx^w$ provides additional information about a role player in the event indicated by the lexical verb. Under the second analysis, there are two event variables at play: one of the events is the event of going, the other is the event of causation. Technically, a third option is also available: the two variables might be posited as lexically independent, but their interpretation might be constrained by providing additional information that they are the same event (perhaps using some grammatical predicate or a con-

straint on indexes using ICONS (Song, 2017)). However, this third option is in many ways like the second because the fact that there are two symbols in the representation allows for the possibility that one may be modified or otherwise reasoned-about independently of the other. For that reason, I set aside the third option for the time being and focus on the first two.

I want to connect these two analyses to some familiar constructions of English. English has a periphrastic causative build around the verb “make” (99a). In the English periphrastic causative, the two variable approach is certainly what we want: in the most natural interpretation of (99a), the adverb *probably* applies to the event of causation and not the embedded state change for *Chris*. In (99b) however, the situation is less clear, I think. Perhaps the morphological transparency of the verb *fell* as a reflection of “cause to fall” allows English speakers to entertain the idea that there are two event structures — one for the causing and a second for the falling. In the third example (99c), however, I find it difficult to entertain the idea that modification of the causation is possible independently of the modification of the bringing. These three examples from English form a cline where on one end (the periphrastic) it is easy to entertain the idea of both an event of causation and an independent event which was caused, but the ability to disassociate the event of causation from the thing which was caused changes from easy to difficult across the three.

- (99) a. Kim [probably] made Chris so sleepy.
 b. Kim [probably] felled the tree.
 c. Kim [probably] brought the child somewhere.

By linguistic analogy to the English constructions in (99), Lushootseed valence-increasing affixes might be seen to be similar to the case of *fell*: derivational morphology creates lexical items which incorporate causative-like structures in a somewhat transparent way. Nevertheless, analysis by analogy to a unrelated languages such as English does no real good at addressing the actual question — whether Lushootseed speakers can entertain the modification of the causative (or applicative) events independently of the lexical predicates they attach to. Instead, these sorts

of questions must be answered experimentally by elicitation from Lushootseed native speakers, or at the worst case, by analogy to elicited results from speakers of closely related languages. But such work is beyond the scope of this thesis. Without this evidence, the conservative course of action is to provide a simpler analysis with fewer variables.

A modified mapping

The DMRS algorithm of Copestake (2009) relies on the characteristic variable property in order to provide a deterministic way to reduce the number of arcs necessary to represent label equality. To recall, multiple EPs may share a label. Analytically, these EPs form a set which will all be collected at a single conjunctive node in a variable scope tree (of a fully scoped-resolved logical form). The naive mapping suggested as a starting point for RMRS to DMRS conversion creates an arc between each pair of such a set. This approach yields a superfluous number of arcs, because the transitive property is inherent in the interpretation of these arcs (i.e. if $\{a, b, c\}$ form a set, minimally two arcs will connect all of the elements, but the algorithm starts by connecting each pair for a total of three arcs). In order to remove superfluous arcs while still maintaining a deterministic mapping, a way to select a canonical element from each set is required. Then, instead of a pair of arcs between each element of the set, only a single arc from the canonical element to each other elements will determine the set minimally and deterministically. The Characteristic Variable Property provides the requisite criterion for selecting a canonical reference. Because my proposed ARG0-sharing representation flouts the Characteristic Variable Property, the mapping algorithm of Copestake (2009) cannot be strictly applied as written.

Modifying the algorithm to accommodate the Lushootseed structures, however, is trivial. The key to making the change is to recognize that the EPs contributed by the valence-increasing affixes are not full predicates in that (by hypothesis, see above) they do not provide any information which should be picked up and modified by any other elements of the construction. Instead, they merely provide further information about the event structure posited by the lexical verb. In this way, they are akin to the pseudo predicates of RMRS. They are “subatomic”, in Parsons’ lingo. I might restate the Characteristic Variable Principle, then, as in (100).

- (100) Let a proper predicate be an Elementary Predication which is neither a generalized quantifier nor a pseudo-predicate, then:

In a well-formed MRS, each variable is the ARG0 of one and only-one proper predicate.

In terms of implementation, there is more than one way to accommodate such elements in the DMRS mapping algorithm. The method I prefer is based on the recognition of predicates contributed by the valence increasing affixes as equivalent to the RMRS pseudo-predicates. This treatment would incorporate the pseudo-predicates into the variable graph, alongside the ARG1s. The idea here is that in the end, what I wish to represent in the two predicates which share an ARG0 in (78) is akin to (101), the *_txw_x_ecs_rel* provides another role player for the event, but does so in neo-Davidsonian way.

$$(101) \left[\begin{array}{l} _?ux^w_go_v_rel \\ ARG0 \ e2 \\ ARG1 \ x4 \\ ECS \ \ x6 \end{array} \right]$$

Preprocessing along these lines would work with no further changes to any libraries. As long as the collection of sub-atomic predicates is listed alongside the other pseudo predicates labels (ARG1_REL, ARG2_REL, etc), the DMRS mapping algorithm requires no further accommodation. An alternative implementation would be to extend the special case already staked out for EPs which represent generalized quantifiers to include the pseudo-predicates.

7.4 Results

The results of the analysis against the testsuite are summarized in Table 7.3. The analysis discussed above provides 100% coverage for the 237 sentences of the testsuite. I reviewed the details of the semantics for each analysis by hand. The syntax and semantics assigned to each item of the grammar is listed in Appendix B.

‘lushootseedmrs/2019-06-01_21:31:16_utc/val-incr/19-06-01/lkb’ Coverage Profile							
Aggregate	total items	positive items	word string	lexical items	distinct analyses	total results	overall coverage
	#	#	ϕ	ϕ	ϕ	#	%
$5 \leq i\text{-length} < 10$	94	94	5.88	6.85	1.00	94	100.0
$0 \leq i\text{-length} < 5$	143	143	2.96	3.17	1.03	143	100.0
Total	237	237	4.12	4.63	1.02	237	100.0

(generated by [incr tsdb()] at 1-jun-2019 (15:12 h))

Table 7.3: Summary of Grammar Coverage for Valence-Increasing Morphology Testsuite (237 sentences from the pedagogical grammar of LR1)

One item of note which emerged from the results was that the prediction of an ambiguity which arises from an interaction of *-yi-t* marking, passivization, and argument dropping. There are four sentences in LR1 Lesson Eight which exhibit this pattern. These are presented in (102) with Hess’ translations.

- (102) a. *ʔuk^wadyitəb* *ʔə tsi luλ’*
ʔu-k^wəda-yi-t-b *ʔə tsi luλ’*
 PFV-take-DAT-ICS-PASS OBL DET.F old.person
 ‘The old woman took [something from someone].’
- b. *ʔuʔabyitəb* *ʔə ti čačas*
ʔu-ʔab-yi-t-b *ʔə ti čačas*
 PFV-extend-DAT-ICS-PASS OBL DET child
 ‘The boy gave [something to someone].’
- c. *ʔuʔuχ^wtx^wyitəb* *ʔə ti luλ’*
ʔu-ʔuχ^w-tx^w-yi-t-b *ʔə ti luλ’*
 PFV-go-ECS-DAT-ICS-PASS OBL DET old.person
 ‘The old man took [something/someone] somewhere for someone.’

- d. *ʔuʔəy̌dx^wyitəb* *ʔə tsi luλ'*
ʔu-ʔəy̌-dx^w-yi-t-b *ʔə tsi luλ'*
 PFV-find-DC-DAT-ICS-PASS OBL DET.F old.person

‘The old woman found [something for someone].’

LR1: Lesson Eight

Recall the example (93), reproduced below as (103a), which shows that the base form without passivization has the argument structure where the subject is associated with the causation (the ICS-role); the first complement (the direct argument) is associated with the DATIVE- (or *yid-*) role; and the oblique is the argument associated closely with the intransitive base: the argument of *ʔab* in the (93) examples (the thing which is given). Hess also provides (103b) which shows how passivization interacts with the three argument construction built by *-yi-t*. In the passivized sentence (103b), the DATIVE argument has been promoted to the subject position, and the former subject (the causer, or ICS-role) is now marked as an oblique.

- (103) a. *ʔuʔabyid* *čəd ti čačas ʔə ti sq^wəbay?*
ʔu-ʔab-yi-t *čəd ti čačas ʔə ti sq^wəbay?*
 PFV-extend-DAT-ICS 1SG.SBJ DET child OBL DET dog

‘I gave the dog to the boy.’

- b. *ʔuʔabyidəb* *čəd ʔə ti čačas ʔə ti sq^wəbay?*
ʔu-ʔab-yi-t-b *čəd ʔə ti čačas ʔə ti sq^wəbay?*
 PFV-extend-DAT-ICS-PASS 1SG.SBJ OBL DET child OBL DET dog

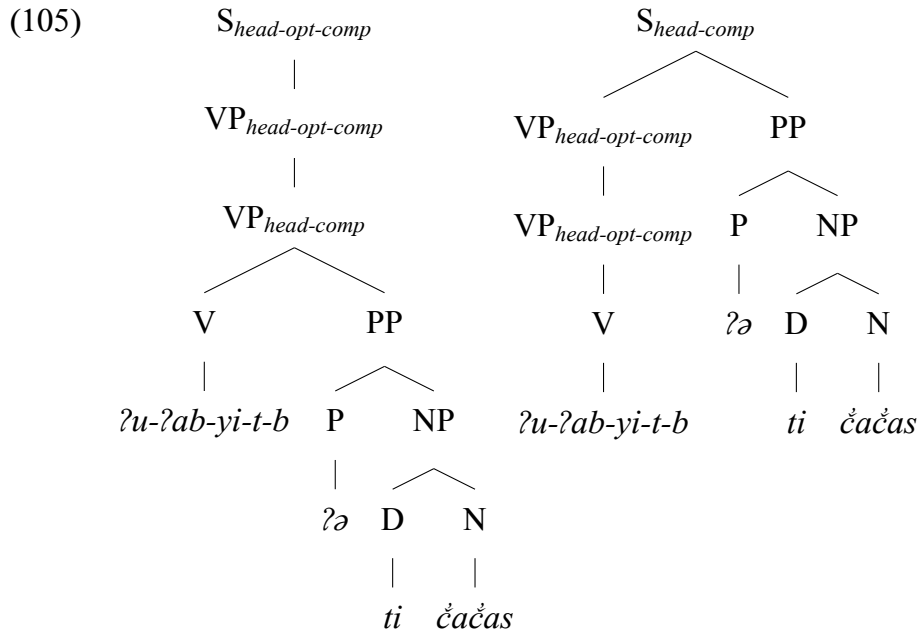
‘The boy gave me the dog.’

LR1: Lesson Eight

In the unpassivized examples with a single argument, the presence or absence of oblique marking distinguishes which argument is missing (see (65) compared to (65b), above), but in the case of a passive, with two oblique arguments, this criterion is unavailable. Because Lushootseed allows the dropping of third person arguments, my grammar predicts a second reading for each

of the four examples of (102) where dropped argument is not the argument of the intransitive base (as it is in the readings given in all four of Hess' glosses), but is the causer, or ics-role. The ambiguity is possible due to the fact that after *yi-d* marking and passivization, both the verb's lexical argument and the ics argument must be headed by the oblique-marking preposition. This, along with the aforementioned argument dropping, allows the two readings. My hypothesized alternative readings are given in free translation lines of (104); I have indicated the linear position of the dropped argument phrase with an ϵ . The two readings correspond to the two syntactic structures illustrated in (105). My hypothesized readings are associated with the structure to the left. Hess' readings are associated with the tree on the right.

- (104) a. *ʔuk^wədyitəb* ϵ ʔə *tsi* *luʔ*.
 ʔu-k^wəda-yi-t-b ϵ ʔə *tsi* *luʔ*
 PFV-take-DAT-ICS-PASS ϵ OBL DET.F old.person
 ‘The old woman was taken [from/for someone by someone].’
- b. *ʔuʔabyitəb* ϵ ʔə *ti* *čačas*.
 ʔu-ʔab-yi-t-b ϵ ʔə *ti* *čačas*
 PFV-extend-DAT-ICS-PASS ϵ OBL DET child
 ‘The boy was given [to/for something/someone by someone].’
- c. *ʔuʔuʔ^wtx^wyitəb* ϵ ʔə *ti* *luʔ*.
 ʔu-ʔuʔ^w-tx^w-yi-t-b ϵ ʔə *ti* *luʔ*
 PFV-go-ECS-DAT-ICS-PASS ϵ OBL DET old.person
 ‘The old man was taken [for/from something/someone by someone].’
- d. *ʔuʔəy^wdx^wyitəb* ϵ ʔə *tsi* *luʔ*.
 ʔu-ʔəy^w-dx^w-yi-t-b ϵ ʔə *tsi* *luʔ*
 PFV-find-DC-DAT-ICS-PASS ϵ OBL DET.F old.person
 ‘The old woman was found [for someone by someone].’



One bit of evidence which may weigh against my hypothesis is that Hess' glosses for these putatively ambiguous sentences consistently drop the final argument (the argument of the intransitive base), keeping the causer. Such consistency may be indicative of a syntactic rule. Since I'm unaware of any further contravening evidence, I have kept these readings for the presented and accepted them as positive results with respect to the testsuite.

7.5 Conclusion

The case study presented in this chapter explores the application of linguistic theory to the modeling of Lushootseed valence-increasing morphology, with a specific focus on event-based semantics as represented by MRS. This is of interest because the linguistic evidence which informs the theory of event-based semantics — at least in the foundational papers I have reviewed — is overwhelmingly drawn from English. Even within the HPSG and MRS literature, there has been little consideration of how the particulars of American languages may inform theory and representation. So there is value, hopefully, in examining the application of these theories and representations to Lushootseed structures.

So how did things go? My evaluation is that the phonological transparency of Lushootseed derivational structure — the easily identified monovalent root paired with equally visible valence-increasing affixes coincides neatly with the Parsonian, or neo-Davidsonian approach where semantically, a verb presents a monovalent predicate whose single argument is the event variable and all role-player arguments are represented as conjunctions. The reason I say that it neatly coincides is that the additional semantic conjuncts which specify role players can be associated with the semantic contribution of the derivational morphology. Such an approach is especially congruent with RMRS, where the role-player-adding predications are part of the formalism.

My own representation, however, did not go so far as Parsons calls for in stripping away core arguments from events. I presented a Davidsonian analysis which is in-line with ERS; I didn't strip away all core arguments. My intransitive verbs come with a single internal argument built in. I did this for two reasons — one practical and one theoretical. The first, the practical one, is that staying Davidsonian allowed me to take advantage of the pre-built machinery of the customization system. I was able to specify Lushootseed verbs as “intransitive” and receive types, rules and instances which interact in tested ways out-of-the-box.⁴¹ The second, theoretical, motivation is mainly that it's unclear what further decomposition would buy me in the present case. Since all Lushootseed verb roots (which I have encountered) have at least one role-player argument in their lexically-specified event structures, creating an additional pseudo-predication to represent this dependency brings additional bookkeeping without any clear payoff. I take the same stance for the handful of underlyingly bivalent verbs in the lexicon. My representations specify their verbal semantics as underlyingly three-place (an event-argument and two role-player slots). Recall that RMRS was built in order to allow maximum flexibility in argument structures so that the output of shallow analysis system such as a part-of-speech tagger could form the input to a deep semantic grammar with full argument indexing. My own use case

⁴¹I did expend some initial effort on setting up the main prerequisite for a “pure” RMRS version of the grammar. This attempt was abortive, unfortunately, as the amount of work required to port the semantic composition mechanisms of the Matrix core grammar from ERS to RMRS proved to be greater than the scope of work I could allocate for this dissertation project.

requires more flexibility than standard ERS provides, but I do not require RMRS flexibility in its full generality.

When it comes to the question of whether to use the mnemonic labels *AGENT*, *PATIENT*, etc. in predicate structures, I have opted not to do so. I prefer the obliqueness hierarchy of *ARG0 ...ARGN*. My rationale here is basically the argumentation in Copestake (2009, 4–5). Consistently choosing and organizing such labels across a large lexicon is not a trivial task. In natural language, particular classes of verb alternations partially cross-cut and overlap each other in ways which make a hand-curation of labels hard to maintain.⁴² Then, assuming Dowty’s proto-role hypothesis as the origin-story for such role-labels as an epiphenomenon deriving from clusters of underlying properties, the actual meanings of the labels become crucially implicated with the entailments of each particular verb’s lexical semantics and so their interpretational value becomes partially eroded anyway. For these reasons, I prefer the numeric indices, which may be associated with an *EXPERIENCER* in one elementary predication and a *CAUSER* in another.

My system, then, is influenced in some way by each of the works discussed in early parts of this chapter: I have pulled from Parsons the idea that sub-atomic semantic structures are observable and useful in particular contexts. I have pulled from Dowty the idea that an ordered argument system allows the definition of thematic role types which may be generalized across verb structures and the idea that organizing these generalizations into an opposition between agent-like and patient-like properties may best explain the obliqueness hierarchies that are lexicalized in particular languages as verbal argument structures. Then, representationally, I have been influenced by RMRS (the Parsonian, or neo-Davidsonian version of MRS) but I do not implement a completely neo-Davidsonian approach. I have drawn on both the motivating arguments for ERS as well as the existing tools implements for creating GMRS (the customization system and core grammar).

My system also breaks with ERS tradition in its violation of the strict version of the Characteristic Variable Principle. Adherence to this principle is a prerequisite for several useful

⁴²See for example Levin 1993 for English. No such work exists yet for Lushootseed.

algorithms which take MRS as input. However, when viewed from the perspective of Parsonian semantic analysis, my violation of the principle is constrained to the subatomic space: only pseudo-predicates introduce ARG0 sharing. Thus, I argue that my violation of the principle is principled and easily circumscribable. I proposed a modification to the Characteristic Variable Principle (and the RMRS-to-DMRS algorithm of Copestake 2009) which would allow my semantics to conform to input assumptions. Should these arguments not suffice, I have also added functionality which allows the system to output MRSEs which are compliant with the unmodified Characteristic Variable Principle.

My system accounts for morphological classes in a mechanical way. I have not implemented any predictive schema such as the one which Hess begins to sketch in Hess (1993). Instead, I have observed a handful of morphological co-occurrences and have engineered an approach which allows for easy extension as I run across new data. I believe that long-term, a robust investigation into Lushootseed verb classes and alternations will be required in order to properly group derivational structures in a predictive way.

The analysis I have presented handles the testsuite without any particular troubles. Beyond the particulars of that analysis, I think that the contribution of this chapter lies in the careful consideration of the analytical alternatives which are available. In the next chapter I continue the exploration of semantic variables and their types but in a different context: nouns and non-verbal predicates.

Chapter 8

CASE STUDY II: VERBAL ARGUMENTS AND NON-VERBAL PREDICATES

This chapter looks at the representation of non-verbal predicates in Lushootseed. In Lushootseed, nouns can appear in predicate position (heading a clause) or in an argument position (heading an NP which is an argument of some other predicate). This case study explores the syntactic and semantic similarities as well as the differences between these two positions. I walk through two implemented systems which roughly correspond to two proposals in the Salishan literature and explore the tradeoffs between simplifying the syntactic analysis at the cost of what is potentially a spurious over-articulation of the semantic space.

The roadmap to this chapter is as follows. I first review the general topic of Salishan nouns. I start with whether and how they are distinguishable from verbs and other Lushootseed parts of speech. This discussion precedes a general presentation of the syntax surrounding non-verbal predicates. As in previous chapters, I present a testsuite which illustrates the relevant phenomena followed by some particulars of the Grammar Matrix customization system which bear on the options for implementation. After the testsuite, as in previous chapters, I present some background information which will bear upon the analysis. In this case, I introduce the Lushootseed general prefixes as these affixes turn out to play a key role in comparing the two systems I will present. Secondly, I introduce a component of the Grammar Matrix customization system which provides some out-of-the-box functionality for capturing the semantic content of the general prefixes.

After that, I present the two implemented approaches and discuss their coverage of the testsuite. The first system is based around the idea of treating nouns predicatively in all syntactic contexts, unifying them with verbs in that respect, and allowing Lushootseed determiners to

expose a nominal index from predicative phrases. This approach is very similar to the proposal in Kinkade (1983) which Beck (2013) argues against. The main difficulty with this approach is that it results in a proliferation of variables and predications in the semantic space which seem to be of questionable value. The second approach limits the application of predicative rules that apply to non-verbal predicates to specific contexts, but this analysis leads to new complications regarding the meaning of a predicative noun. I then look to some further data in the Lushootseed corpus in order to shed light on whether the practice of limiting predication contexts is really viable. This is where the Lushootseed general prefixes are of diagnostic use. I focus especially on the past tense marker in examples where it appears on nouns.

Finally, I summarize the systems and their tradeoffs. In contrast to the word-order discussion of Chapter 6, where I opted for one system over the other, in this case I opt to kick the can down the road, taking advantage of the CLIMB methodology of maintaining multiple approaches to a phenomenon in a configurable metagrammar. I continue evaluating the two systems presented here in the next chapter.

8.1 Nouns and Verbs in Salishan

No topic from Salishan linguistics is more well-known to the general linguistic audience than the noun-verb question: some authors (Kuipers, 1968; Kinkade, 1976; Thompson and Thompson, 1980; Kinkade, 1983) have claimed that Salishan languages lack a distinction between nouns and verbs. Others (Hess and van Eijk, 1985; van Eijk and Hess, 1986; Beck, 2013) have replied that these languages *do* indeed evince a noun-verb distinction and have listed the particular grammatical contrasts which bear this out. As far as I can tell, the disagreement is due to the fact that if there are nouns and verbs in Salishan languages, they overlap with respect to the syntactic positions in which they can be found. So some authors have sought analytical pathways which explain the data without reference to these categories. Because the resolution of this question bears directly on the available analyses, I first recap the historical arguments against the distinction, and then try to show how such arguments seem difficult to apply to the practical task of the implemented grammar writer.

The paper which is usually cited first in this realm is Kuipers (1968) “The categories Verb-Noun and Transitive-Intransitive in English and Squamish”¹. If I understand the argumentation correctly, Kuipers believes that the distinctions between the intransitive and transitive paradigms in lexical items are sufficient to account for the grammatical distinctions observable in Squamish. However, to my reading, the paper seems to rely overly on a requirement of semantic translation equivalence between the two languages. Kuipers dismisses the overt morphological marking which appears on the forms which have been traditionally called nouns by observing that not all Squamish words which translate to nouns in English bear this prefix. He also cites the converse, that not all English nouns are translated into Squamish words which bear this so-called nominal morphology (p. 612). While the details of Kuipers’ formalization are interesting, much of the argumentation relies overly on this type of linguistic grounding: Starting from a hypothesis that the difference between nouns and verbs in English is fully characterized by the fact that only one of these two classes can be predicative, Kuipers observes that because Squamish nouns *are* predicative, it follows that there is no distinction between these categories.

The next set of papers which argue against a noun-verb distinction are those by Dale Kinkade (1976; 1983). Kinkade (1976) is mostly descriptive, but does set the stage for the more theoretical paper (Kinkade, 1983), of which the noun-verb distinction is the primary topic. My reading of Kinkade’s arguments are that they hinge on theoretical notions which abstract away from surface phenomena in such a way that both the empirical falsifiability of the hypothesis and its application to descriptive or empirically based modelling becomes difficult to ascertain. For example, Kinkade sweeps away the morphological evidence by ascribing it to a reflection of semantics:

One criterion that is often used to distinguish word classes is the restriction placed on occurrence of various affixes. I see no basis for claiming that there are limitations on such co-occurrences in Salish that are morphological or syntactic; certainly there are limitations having to do with semantic incompatibility, as in any language. (Kinkade, 1983, 27)

¹Skwxwú7mesh [squ], anglicized as “Squamish” is a Central Salish language of British Columbia. See Chapter 2 for some notes on the distribution and philogenetic subgroupings of Salishan languages.

I think two items are noteworthy here: The first is that Kinkade's argumentation provides a curious inverse to that of Kuipers. For Kinkade, morphological distinctions are just a reflection of (lexical) semantics, and therefore uninteresting. On the other hand, Kuipers' claim is that semantics must drive the interesting distinction, and because morphological marking is flexible in some semantic contexts when English provides the yardstick, the morphology must not be of interest. The second item to note, and this one is all the more powerful to me insofar as it directly relates to my concerns as a grammar engineer, is that the descriptive papers which share the theoretical concerns of Kuipers and Kinkade continue to refer to noun phrases (albeit with some hedging), presumably because the utility of this notion in surface-oriented description outweighs any theoretical purity which may also be of interest:

The term "noun-phrase-like" is used here because of an important Salish characteristic: when the familiar noun/verb opposition of commonly studied languages is projected on these structures, the systems are distorted in important ways. This problem is explored in M. Dale Kinkade, "Evidence Against the Universality of 'Noun' and 'Verb': Salish" (University of British Columbia, 1978)

(Thompson and Thompson, 1980, 31)

On the other side of the debate, the authors who argue for a noun-verb distinction cite these same types of morphological data that Kinkade and others discount: that Lillooet and Lushootseed only allow aspect and transitive morphology on verbs; they allow possessive morphology only on nouns. Hess and van Eijk (1985) acknowledge and address the theoretical argumentation of Kinkade while motivating, from a descriptive and theoretical perspective, the noun-verb distinction in concrete examples for Lillooet and Lushootseed. The analysis in Beck (2013) refines the problem under a model in which parts of speech may be associated with "criterial" syntactic positions and a neutralization or lack of syntactic distinction between parts-of-speech is defined in terms of structural markedness which may (or may not) be present when the words of one category appear in syntactic positions which are criterial for another category. Specifically, in terms of Lushootseed, Beck agrees with previous work that Lushootseed (and other Salish lan-

guages) display a neutralization in predicate position: nouns and other parts of speech appear as predicates without any additional morphological marking. However, he disagrees with preceding authors that there is corresponding neutralization when verbs appear in argument position (the position criterial for nouns). Beck's examples of this lack of full neutralization are drawn from the fact that when bare verbs appear in argument position, they are interpreted as subject-centered headless relative clauses. As I understand the argument, because some morphological process must be creating this relativization, there is a structural difference between bare verbs in argument position and nouns in argument position, and therefore there is no full neutralization between nouns and verbs at the syntactic level, in contra previous works. Beck's analysis is interesting because he frames the question of what it means to have a syntactic neutralization in concrete analytical terms. Since I am building an analytical implementation, this type of framing is one which I may be able to answer to — it is a question I can return to more fully after presenting the data and the implementations.

For my purposes, the question of whether or not Lushootseed exhibits a noun-verb contrast is a settled one: the Lushootseed data clearly exhibits morphological phenomena which are perspicaciously described along such lines. Then, the interesting thing for me is to follow this fact with an exploration of the particulars of the syntax and semantics which these elements exhibit: exactly where they overlap and where they differ is part of the point of this chapter. After implementing a system to capture the testsuite, I can answer as to whether a syntactic neutralization of these categorizes exists in my grammar in terms of the explicit analytical criteria laid out in Beck (2013).²

²More broadly, on the controversy itself, I think that Haspelmath (2012) is correct in framing the differing answers to the noun-verb question of Salish as reflective of a broader pendulum swing away from the universalist tendencies of the late 20th century in linguistic typology. Haspelmath cites Bossong (1992) for this observation.

In context, these universalist tendencies were themselves a response to the particularism of the early 20th century, which can be seen as a response to the universalism of the Rationalist ideas of the Enlightenment period (e.g. Chomsky 1966).

8.2 *Data*

In this section I present a selection of data which forms my Lushootseed testsuite for this case-study. Following that, I offer some background on particular aspects of the Grammar Matrix customization system which bear on the discussion that follows this section.

8.2.1 *Testsuite*

In this subsection I will provide some illustrative examples of the phenomena I wish to cover. As in previous chapters, the goal is to build up a testsuite which can be used for concrete analysis of candidate approaches. Because the main thesis of Hess and van Eijk (1985) is to demonstrate that there are categorical differences between noun and verb in Lushootseed, the example sentences from that paper provide a handy starting point for building a testsuite. I begin by walking through the examples from that paper which I included in the testsuite.

Probably the most iconic of the examples illustrating that nouns are equally viable as the main predicate position of a sentence is one such as (106a), in which a noun takes a subject marker.³ Third-person subjects are not marked, so (106b) is also a full sentence.

(106) a. *stubš čəd*
 man 1SG.SBJ
 ‘I am a man.’

b. *stubš*
 man
 ‘He is a man.’

(Hess and van Eijk, 1985)

As mentioned in the introduction, Lushootseed surface syntax is happy to place verbal elements in argument position, just as nominal elements may be placed in predicate position. The

³Recall from Chapter 6 that I treat these subject markers as auxiliary verbs.

two examples of (107) show this. While apparently truth-conditionally equivalent, that these two examples do ostensibly differ in their discourse properties. The particulars of Lushootseed discourse and information structure lie outside of the scope of my current investigation.

(107) a. *ʔuχ^w ti sbiaw*
go DET coyote

‘The coyote goes.’

b. *sbiaw ti ʔuχ^w*
coyote DET go

‘It is the coyote that goes.’

(Hess and van Eijk, 1985)

The next pair of sentences (108) is like those in (107) in that it shows a typical Verb-Subject construction and its clefted counterpart. The difference is that in (108), the verb bears transitive and object marking. Because valence-increasing and object-marking morphology can only appear on verbal bases (Hess and van Eijk 1985; Beck n.d.), this example shows that even in (108b), *k^wax^wac* is verbal.

(108) a. *k^wax^wac ti stubš*
k^wax^wa-t-s ti stubš
help-ICS-1SG.OBJ DET man

‘The man helped me.’

b. *stubš ti k^wax^wac*
stubš ti k^wax^wa-t-s
man DET help-ICS-1SG.OBJ

‘It is a man who helped me.’

(Hess and van Eijk, 1985)

Similarly to how (108) shows that the transitivizing morphology which may appear on verbs is orthogonal to their ability to appear in canonically nominal positions, the examples in (109)

show the converse. Possessive marking only appears on nouns (Hess and van Eijk, 1985; Beck, n.d.), so *dqsi* and *adqsi* must be nominal, even when these lexical items are functioning as the main predicate.

- (109) a. *dqsi* *čəx^w*
 d-qsi čəx^w
 1SG.POSS-uncle 2SG.SBJV

‘You are my uncle.’

- b. *adqsi* *čəd*
 ad-qsi čəd
 2SG.POSS-uncle 1SG.SBJV

‘I am your uncle.’

(van Eijk and Hess, 1986)

I also pulled data from Lushootseed Reader Volume I (Hess, 1995, LR1), as in previous chapters, to construct a targeted dataset. The examples in the beginning of §16.1 of (LR1) specifically treat the fact that some Lushootseed sentences have no verbs at all. So I included these examples in my testsuite as well. These sentences will end up forming crucial examples for this case study because of the tense and mood markers which they also display (110).

- (110) a. *ʔubəščəb ti?iʔ*
 ʔu=bəščəb ti?iʔ
 IRR=mink DIST.DEM

‘That one will become a mink.’

- b. *tusi?ab ti tudčístx^w*
 tu=si?ab ti tu=d-sčístx^w
 PST=person.of.rank DET PST=1SG.POSS-husband

‘My former husband was a man of rank.’

LR1 p.81

Lushootseed has special forms of the personal pronouns which have the syntactic distribution of nouns. These full pronouns are rare because in most contexts, speakers prefer to mark first- or second-person arguments with subject- and object-markers discussed in Chapter 6. The example in (111) shows that the full-noun version of these pronouns can also be predicative.

- (111) *?əca ti tu=diʔqs*
?əca ti tu=diʔ•qs
 1SG DET PST=other.side•point

‘I am the one who was on the other side of the point.’

Lushootseed predicative position is not restricted to nouns and verbs. Adjectives may also be placed into predicative position (112). This item is included in the testsuite and serves to illustrate that the topic at hand is really not just about nouns and verbs, but about non-verbal predicates more generally. Nevertheless, I will tend to talk about nouns since most of the examples under discussion are about nouns.

- (112) *saliʔ tiʔəʔ sqigʷac*
 two DET deer

‘There are two deer.’

LR1 p.81

The demonstrative pronouns may also appear in predicate position, as illustrated in (113).

- (113) a. *tiʔəʔ tə čʔaʔ*
 PROX.DEM DET rock

‘This is the rock.’

- b. *tudiʔ tə dukʷibəʔ*
 DIST.DEM DET changer

‘Way off over there is Changer.’

LR1 p.81

Finally, my initial testsuite pulls in two examples which illustrate the fact the subject-centered relative clause which is formed by a verbal element in an argument phrase may be complex. This can be seen in (114a) where the verb *ʔučalad* (“chased”) along with its complement are both embedded in argument position under the determiner *tiʔəʔ*. The verb combines with its object (*tiʔəʔ sqʷəbayʔ* (“the dog”)) within the determiner phrase and then the referent of the entire phrase is the verb’s subject (the chaser). When speakers want to refer to the non-subject argument of a transitive which is embedded in a determiner phrase, they create a passive in order to get the target argument into subject position (114b).

- (114) a. *wiʷsu tiʔəʔ ʔučalad tiʔəʔ sqʷəbayʔ*
wiʷsu tiʔəʔ ʔu-čala-t tiʔəʔ sqʷəbayʔ
 children DET PFV-chase-ICS DET dog

‘The children are the ones who chased the dog.’

- b. *sqʷəbayʔ ti ʔučalatab ʔə tiʔiʔ wiʷsu*
sqʷəbayʔ ti ʔu-čala-t-b ʔə tiʔiʔ wiʷsu
 dog DET PFV-chase-ICS-PASS DET children

‘The dog is the one the children chased.’

LR1 p.99

At 17 items in total (10 from Hess and van Eijk (1985); van Eijk and Hess (1986), 7 from LR1), this testsuite is small (especially compared to the 237 items which were collected for the testsuite associated with the preceding case study). Nevertheless, it does provide a targeted set of sentences which lay out the data to be modelled. Also, both of the systems presented below in the Chapter are also evaluated against a larger set of sentences drawn from naturally running text. This later evaluation is described in Chapter 9.

Name	Gloss	Form
Additive	ADD	bə=
Subjunctive	SBJ	g ^w ə=
Hab	PAST	ʔu=
Irrealis	IRR	ʔu=
Past	PAST	tu=

Table 8.1: Lushootseed General Prefixes

8.2.2 *Lushootseed General Prefixes and Contentful Morphology in the Grammar Matrix Customization System*

In this section I provide some background which pertains to the past-tense *tu=* and the irrealis *ʔu=* markers which appear in the testsuite. These markers are part of a class of morphology which Hess (1995, 62) calls general prefixes. The generality referred to is the fact that these markers may attach to any particular word class “so long as it functions as the main word in a predicate or complement”. Beck analyses these markers as phrase-level clitics (Beck, n.d.), so the data in the examples and in my testsuite uses the clitic marker (=). I have treated these examples as prefixes in my grammar rather than clitics as a first approximation, not because they oughtn’t be clitics.

The Past and Irrealis markers are in complementary distribution (Hess, 1995, 64). For this reason, I treat them both as values of the TENSE feature in my grammar. The irrealis marker (*ʔu=*) is also referred to by some authors as marking future tense, as opposed to irrealis mood (Bates and Hess, 2001). While I remain agnostic with respect to the proper cross-linguistic categorization of the functions of these markers, formally, Lushootseed has two aspect markers which are in complementary distribution and with which Irrealis cooccurs. This fact, along with the fact that that it does not cooccur with Past (*tu=*) makes a formal analysis as tense convenient. However, I maintain Beck’s gloss as Irrealis for consistency with my data source.

Both the past and irrealis markers show up (on nouns) in examples in the testsuite (110). While they don’t happen to appear in the testsuite on verbs, they are attested there as well. The same is true of the other markers listed above — they don’t actually appear in this testsuite, but

they do appear in the corpus on both nouns and verbs.

In terms of treating these prefixes, the Grammar Matrix (Bender et al., 2010) customization system provides some relevant functionality out-of-the-box. The meanings marked by these elements are for the most part congruent with traditional features of events: tense, aspect and mood. A user of the customization system can supply their own values to be used in tense, aspect, and mood hierarchies (Pouson, 2011). These options are somewhat integrated with the morphotactics system in that once these hierarchies are defined, they be associated (with some limitations) with lexical rules defined in the morphotactics pages. A set of choices which illustrate this specification is presented in Figure 8.1.

When a user of the customization system requests tense, aspect and mood values as shown in Figure 8.1, the output grammar contains a corresponding hierarchy defining these types and the types which define the rules (115). The Matrix core grammar already defines event type indices as bearing tense, aspect and mood values.

```
(115)   ;;; Tense
        past := tense.
        irr := tense.

        ;;; Aspect
        perfective := aspect.
        imperfective := aspect.
        stative := aspect.
        progressive := aspect.
        progressive-state := aspect.
        continuous := aspect.

        ;;; Mood
        hab := mood.
        add := mood.
        sbj := mood.

        irr-lex-rule := proclitics-lex-rule-super &
          [ SYNSEM.LOCAL.CONT.HOOK.INDEX.E.MOOD irr ].

        past-lex-rule := proclitics-lex-rule-super &
          [ SYNSEM.LOCAL.CONT.HOOK.INDEX.E.TENSE past ].
```

```

section=tense-aspect-mood
  tense-definition=build
  tense1_name=past
    tense1_supertype1_name=tense
  tense2_name=irr
    tense2_supertype1_name=tense

  aspect1_name=perfective
    aspect1_supertype1_name=aspect
  aspect2_name=imperfective
    aspect2_supertype1_name=aspect
  aspect3_name=stative
    aspect3_supertype1_name=aspect
  aspect4_name=progressive
    aspect4_supertype1_name=aspect
  aspect5_name=progressive-state
    aspect5_supertype1_name=aspect
  aspect6_name=continuous
    aspect6_supertype1_name=aspect

  mood1_name=hab
    mood1_supertype1_name=mood
  mood2_name=add
    mood2_supertype1_name=mood
  mood3_name=sbj
    mood3_supertype1_name=mood

section=morphology
  lex-pc2_name=proclitics
  lex-pc2_obligatory=off
  lex-pc2_order=prefix
  lex-pc2_inputs=noun, noun-pc1, verb-
    -pc1, verb-pc2, verb-pc3, verb-
    pc4, lex-pc1
  lex-pc2_lrt1_name=hab
  lex-pc2_lrt1_feat1_name=mood
  lex-pc2_lrt1_feat1_value=hab
  lex-pc2_lrt1_feat1_head=verb
  lex-pc2_lrt1_lril_inflecting=yes
  lex-pc2_lrt1_lril_orthʰ=u=

  lex-pc2_lrt2_name=irr
  lex-pc2_lrt2_feat1_name=mood
  lex-pc2_lrt2_feat1_value=irr
  lex-pc2_lrt2_feat1_head=verb
  lex-pc2_lrt2_lril_inflecting=yes
  lex-pc2_lrt2_lril_orthʰ=u=

  lex-pc2_lrt3_name=past
  lex-pc2_lrt3_feat1_name=tense
  lex-pc2_lrt3_feat1_value=past
  lex-pc2_lrt3_feat1_head=verb
  lex-pc2_lrt3_lril_inflecting=yes
  lex-pc2_lrt3_lril_orth=tu=

  lex-pc2_lrt4_name=add
  lex-pc2_lrt4_feat1_name=mood
  lex-pc2_lrt4_feat1_value=add
  lex-pc2_lrt4_feat1_head=verb
  lex-pc2_lrt4_lril_inflecting=yes
  lex-pc2_lrt4_lril_orth=əb=

  lex-pc2_lrt5_name=sbj
  lex-pc2_lrt5_feat1_name=mood
  lex-pc2_lrt5_feat1_value=sbj
  lex-pc2_lrt5_feat1_head=verb
  lex-pc2_lrt5_lril_inflecting=yes
  lex-pc2_lrt5_lril_orth=wəg=

```

Figure 8.1: A set of tense, aspect and mood choices implementing Lushootseed general prefixes.

The customization system also allows users to define their own hierarchies of values for certain features and to refer to them in lexical rules. However, these custom hierarchies, in the default system, don't extend the definition of event-type variables. Instead, the semantic features which can be defined in the customization system are added to the definition of instance-type variables, making them (effectively) appropriate for semantic indices that are usually associated with nominals.⁴

8.2.3 *Predicative nouns in the Matrix-derived grammars*

Nouns which appear in predicate position without a supporting copula are not rare across the world's languages (Stassen, 2013). While the Grammar Matrix customization system does not provide out-of-the-box support for such, there is a recommended approach for extending a Matrix-based grammar to analyze predicative nouns which is published as part of the course-materials for Emily Bender's grammar engineering course at the University of Washington.⁵

The recommended approach is to create a non-branching phrase-structure rule which takes a daughter of `HEAD noun` and which yields a verb phrase. Semantically, such a rule posits an event-type variable to act as the `INDEX` of the new phrase as well as the `ARG0` of a newly-positing copula predicate. The example provided by the course materials also provides a quantifier to semantically bind the daughter's `INDEX` variable. The example rule in the course documentation from 2016 is provided in (116).

⁴§3.5.2 introduces the hierarchy around variable-types.

⁵<https://courses.washington.edu/ling567/2016/lab6.html>; Accessed 2019-07-31T00:22+00:00.

```

(116) n-bar-predicate-rule := unary-phrase & nocoord &
      [ SYNSEM.LOCAL.CAT [ HEAD verb,
                          VAL [ COMPS < >,
                                SUBJ < [ LOCAL [ CONT.HOOK.INDEX #arg1,
                                                CAT [ HEAD noun,
                                                    VAL.SPR < > ] ] ] > ] ],
      C-CONT [ HOOK [ LTOP #ltop,
                    INDEX #index,
                    XARG #arg1 ],
      RELS <! arg12-ev-relation &
            [ PRED "_be_v_id_rel",
              LBL #ltop,
              ARG0 #index,
              ARG1 #arg1,
              ARG2 #arg2 ],
            quant-relation &
            [ PRED "exist_q_rel",
              ARG0 #arg2,
              RSTR #harg ] !>,
      HCONS <! qeq & [ HARG #harg, LARG #larg ] !> ],
      ARGS < [ SYNSEM.LOCAL [ CAT [ HEAD noun,
                                   VAL.SPR cons ],
                                CONT.HOOK [ INDEX #arg2,
                                             LTOP #larg ]]] > ].

```

This approach makes two linguistic predictions, that I see. First, it suggests that after the application of such a rule, these predicative nouns behave like verbs. They are literally made into verbs by this rule so any further rules which require a verbal daughter may apply. Secondly, the rule as I have written it here predicts that no index-binding elements — such as determiners in the traditional approach — are allowed on predicative nouns (the course notes I referenced above point out an alternative version which leaves out the quantifier binding aspects of this version). These predictions are explored in the next section where I begin to flesh out two implementable approaches to the data which build on the background presented here.

8.3 Analysis

In this section, I present two approaches to the data and compare the tradeoffs between them. My starting point for the discussion below is a grammar which has been augmented with a rule very similar to the one in (116). The main difference is that I have used a lexical, rather than a syntactic rule, to derive the predicative noun from the base one. I believe that doing this work in the lexicon is preferable to the syntax, at least for Lushootseed, as it fits with the idea that the non-controversial differences between nouns and verbs in Lushootseed are the morphological ones — that is, it keeps nominal and verbal syntax unchanged and allows me to test out how different versions of the lexical rule interact with the shared syntactic rules.

For organizational purposes, I created a supertype which defines the structure that is shared by all of the non-verbal predicate types. This is shown in (117). The discussion below presents specific rules which inherit from the rule in (117).

```
(117) non-verbal-pred-lex-rule-super := lex-rule &
      [ SYNSEM.LOCAL.CAT
        [ HEAD [ AUX -,
                PRD +,
                MOD < > ],
          VAL [ COMPS < >,
                SPEC < >,
                SUBJ < [ LOCAL [ CONT.HOOK.INDEX ref-ind & #arg1,
                                CAT [ HEAD noun,
                                      VAL [ SPR < >,
                                            SUBJ < >,
                                            SPEC < >,
                                            COMPS < > ] ] ] ] > ] ],
          C-CONT.HOOK.XARG #arg1 ].
```

Then, for adjectives or adverbs, which provide their own predication whose argument is lexically linked into their on their MOD list, my system provides a rule which connects the element which would have been modified to SUBJ list, creating a subject dependency with properly linked semantics. My rule is shown in (118).

- (118) a-pred-lex-rule := non-verbal-pred-lex-rule-super &
 [SYNSEM.LOCAL.CAT.VAL.SUBJ < [LOCAL.CONT.HOOK.INDEX #x] > ,
 C-CONT [HOOK [INDEX #e,
 LTOP #h] ,
 RELS <! !>] ,
 DTR.SYNSEM.LOCAL [CAT.HEAD +jr &
 [MOD < [LOCAL.CONT.HOOK.INDEX #x] >] ,
 CONT.HOOK [INDEX #e,
 LTOP #h]]] ."""

Nouns are a bit more complicated for at least two reasons. The first is that in the standard approach to predicative nouns suggested for Matrix-based grammars, the predicative noun must provide a copula relation. The copula relation provides an event-like variable which is appropriate for features associated with predicates like tense/aspect/mood. The second complication regards variable binding. The noun provides an individual-type variable and well-formed semantics must provide a quantifier and scope under which assignment may take place. For non-predicative nouns, the approach thusfar has been that determiners provide the necessary quantifiers, yet the sentence-initial predicative nouns in the testsuite do not bear determiners. So getting the variable-binding correct in these constructions is a second matter which must be addressed.

The rule in (119) adds the copula relation and connects its daughter's INDEX to the ARG2. It enforces a non-null SUBJ list on its daughter and links the subject dependency's INDEX to the ARG1 of the copula relation using the XARG feature. It handles the case of sentence-initial predicative nouns without an explicit determiner by providing a generic *exist* quantifier and binding off it's own index variable in the standard way using a handle constraint.

```
(119) n-pred-lex-rule := non-verbal-pred-lex-rule-super &
      [ C-CONT [ HOOK [ INDEX #index,
                    XARG #arg1,
                    LTOP #ltop ],
          RELS <! arg12-ev-relation &
                [ PRED "_copula_v_id_rel",
                  LBL #ltop,
                  ARG0 #index,
                  ARG1 #arg1,
                  ARG2 #arg2 ],
                quant-relation &
                [ PRED "exist_q_rel",
                  ARG0 #arg2,
                  RSTR #harg ] !>,
          HCONS <! qeq & [ HARG #harg, LARG #larg ] !> ],
      DTR [ SYNSEM.LOCAL [ CAT.VAL.SPR cons,
                          CONT.HOOK [ INDEX #arg2,
                                      LTOP #larg ] ] ] ].
```

These rules for the starting point for a discussion of the two systems I considered. Depending on which system (of the two I describe below) is employed, there are a couple of other constraints which are added to the rules above in order to complete the implementation of non-verbal predicates. In the next two sections I describe two these two approaches as System I: Everything is a predicate and System II: Limited predication contexts. These two systems are (respectively) linked to the arguments for and against the idea that there is no syntactic distinction between nouns and verbs in Lushootseed.

In the discussion of these systems, I will often refer to an embedded context. The type of embedding I'm referring to is appearing in a phrase headed by a determiner. I contrast this with predicative position, which is generally sentence-initial.

8.3.1 *System I: Everything is a predicate*

The fact that nouns and verbs (or verb phrases) can ostensibly appear in the same syntactic contexts leads, naturally, to thinking about how they may be modelled as equivalent structures at the syntactic level. As I began looking at extending the Grammar Matrix determiner types

to accommodate a verb-headed argument, I considered that once I have a predicative-noun rule which can create verbs from nouns (even, without further constraint, from nouns which are embedded under determiners), a natural extension of the grammar to cover embedded verbs would be to allow determiners to take these verbal elements and to publish an index referring to the subject.

Such an approach unifies embedded verb phrases with embedded nominals and is very similar to the system which is proposed in Kinkade 1983 for Salishan generally (and which is forcefully argued against in Beck 2013). Under this approach, the determiners really function more like subject-marking relative pronouns in a free-relative construction, by attaching to a phrase which lacks a subject and exposing the INDEX of the predicate's subject, rather than the INDEX of the predicate itself.

I have implemented this system in my metagrammar by adding a new determiner type to the grammar (120). This type requires a verbal head, this is indicated by the HEAD value of the single element on the SPEC list (the feature SPEC is where determiner types create selection criteria for the heads they can attach to). It also takes the referential index of its SPEC's subject, using the XARG feature for access, and assigns this index into its own feature structure. This type is very similar to the *basic-determiner-lex* provided by the Matrix core grammar, the only real differences being the SPEC value selecting for a verb instead of a noun and binding the verb's subject's INDEX as opposed to a noun's INDEX.⁶

⁶As pointed out by Bender (pc), this rule actually breaks the Characteristic Variable Principle (CVP) discussed in Chapter 7 because there is no predicate whose ARG0 is value tagged #x in this rule. This could be remedied by having the rule contribute a grammatical predicate, a *nominalization_rel*, whose only purpose would be to host the posited variable as ARG0, thereby formally satisfying the CVP.

```

(120) subj-raising-determiner-lex := norm-sem-lex-item & non-mod-lex-item &
      [ SYNSEM
        [ LOCAL [ CAT [ HEAD det,
                      VAL [ SPEC
                          < [ LOCAL [ CAT.HEAD verb,
                                      CONT.HOOK [ XARG #x,
                                                  LTOP #larg ] ] ] >,
                          SUBJ < >,
                          COMPS < > ] ],
          CONT [ HCONS <! qeq & [ HARG #harg,
                                  LARG #larg ] !>,
                RELS <! relation !>,
                HOOK.INDEX #x ] ],
          LKEYS.KEYREL quant-relation &
            [ ARG0 #x,
              RSTR #harg ] ] ] ].

```

A phrase built using this determiner type should be of HEAD type *noun*, since (1) it exposes an instance-type INDEX and (2) I want it to be able to function as an argument to higher predicates. The determiner, however, is not the head of a typical head-spec phrase. Therefore, in order to get the head-type correct on the output, I took a constructional approach to this, and did the rest of the work for this approach by creating a special phrase-structure rule-type which combines this determiner with the verb and yields a saturated noun — an NP. The INDEX of this phrase, as I mentioned above, is identified with the INDEX of the verb's subject (121). This rule is tied into the head-marker types in order to approximate headedness without actually requiring token-identity of the HEAD values, since the rule wants to change the head-value to NOUN.

```

(121) subj-raising-spec-head-phrase := marker-initial-phrase &
      [ SYNSEM [ LOCAL.CAT [ HEAD noun & [ MOD < > ],
                          VAL [ SUBJ < >,
                                COMPS < > & #spcomps,
                                SPR < > & #spr,
                                SPEC < > & #spec ],
                          POSTHEAD #ph,
                          HC-LIGHT - ],
                          MODIFIED #modif ],
      MARKER-DTR
      [ SYNSEM #synsem &
        [ LOCAL [ CAT [ VAL [ SPEC
                          < [ LOCAL [ CAT [ HEAD #head,
                                      VAL.COMPS #comps ],
                                      CONT.HOOK #hdhook ],
                          MODIFIED #hmodif ] >,
                          COMPS #spcomps ] ],
          CONT.HOOK #hook ],
          MODIFIED #modif ] ],
      NON-MARKER-DTR
      [ SYNSEM [ LOCAL [ CAT [ HEAD #head,
                          VAL [ SUBJ cons,
                                COMPS olist & #comps,
                                SPR < #synsem & canonical-synsem
                                  . #spr >,
                                SPEC #spec ],
                          POSTHEAD #ph ],
                          CONT.HOOK #hdhook ],
          MODIFIED #hmodif ] ],
      C-CONT [ HOOK #hook,
              RELS <! !>,
              HCONS <! !>,
              ICONS <! !> ] ].

```

This system provides one further constraint on the two lexical rules of (117) and (118) which create predicative nouns and adjectives from their base versions: it adds the constraint that the output of the rule is HEAD *verb*. Those rules, along with the special subject-raising determiner construction just described, when compiled into a grammar, are enough to get respectable coverage over the 17 item testsuite — only the last two sentences, those of (114), in which the embedded verb has an overt complement, fail to parse.

For the simple examples in (106), the grammar output is correspondingly straightforward. The predicative-noun lexical rule applies to *stubš*, creating an intransitive verb. The subject marking auxiliary attaches to the predicative noun the same way it attaches to verbs (122a).⁷ The MRS assigned by the grammar to this example is shown in (122b).

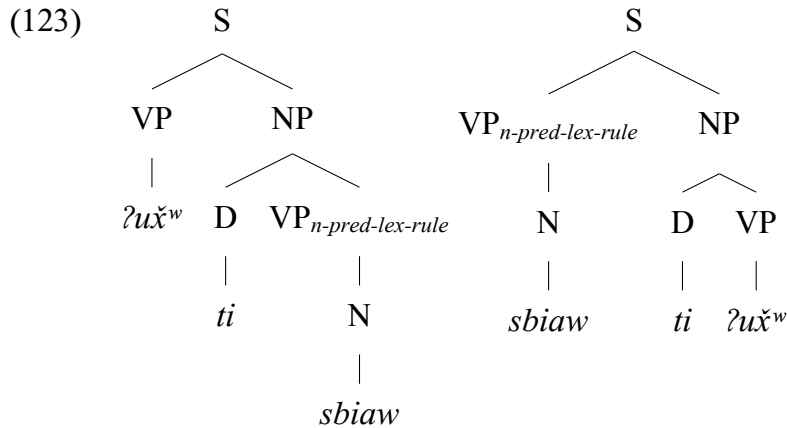
- (122) a.
- ```

graph TD
 S --> VP["VPn-pred-lex-rule"]
 S --> Aux
 VP --> N
 N --> stubš
 Aux --> čəd

```
- b.
- |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |            |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|
| $\langle h_1, e_2,$<br>$h_3: \text{"_stubš\_n\_man\_rel"}(\text{ARG0: } x_4\{\text{PNG.PER } 3RD\}),$<br>$h_5: \text{"_copula\_v\_id\_rel"}(\text{ARG0: } e_2$<br><div style="margin-left: 100px;"><math>\text{ARG1: } x_6\{\text{PNG.PER } 1ST, \text{PNG.NUM } SG\}</math></div> <div style="margin-left: 100px;"><math>\text{ARG2: } x_4)</math></div> $h_7: \text{"exist\_q\_rel"}(\text{ARG0: } x_4, \text{RSTR: } h_8, \text{BODY: } h_9)$<br>$\{ h_8 =_q h_3 \} \rangle$ | $,$<br>$,$ |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|

Under this system, the syntax of the (a) sentences looks very similar to the syntax of the (b) sentences (107) and (108). The two trees for (107) are presented in (123). The similarity I refer to comes from the fact that because determiners are defined to take verb-like elements, in the (a) examples, the only way to create an NP is for the predicative-noun rule to fire. Then, in the (b) examples, where a verb and determiner phrase create an expression referring to the subject of the verb, the same specialized NP construction is used. The result is that subject of the main lexical predication of *ʔuχʷ* (“go”), in one case, or the copula relation in the other case, provides the INDEX which is picked up as the ARG1 of the main predicate of the clause.

<sup>7</sup>See Chapter 6.



The semantics is correspondingly similar: both sentences contain the same collection of elementary predications with the exact same links between them (124). The only difference is which of the two event-type predications is *topmost INDEX* of the MRS. In the verb-initial sentence (107a), it is the *\_?uχʷ\_v\_go\_rel*. In the noun-initial example (107b), it is the copula relation.

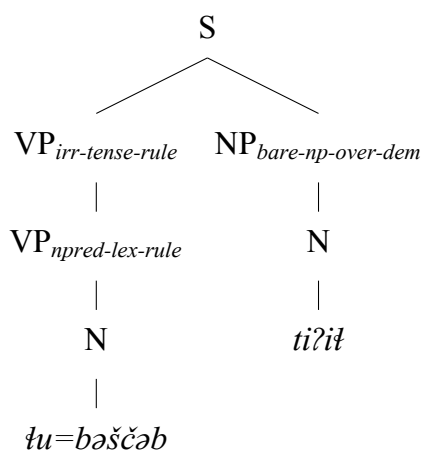
(124)  $\langle h_1, e_2\{\text{SF PROP-OR-QUES, E.ASPECT IMPERFECTIVE}\},$   
 $\left. \begin{array}{l} h_3: \text{"_?u}\chi^w\_v\_go\_rel"(ARG0: e_2, ARG1: x_4), \\ h_5: \text{"_sbiaw\_n\_coyote\_rel"(ARG0: x_6\{PNG.PER 3RD\}), \\ h_7: \text{"_copula\_v\_id\_rel"(ARG0: e_8, ARG1: x_4, ARG2: x_6), \\ h_9: \text{"exist\_q\_rel"(ARG0: x_6, RSTR: h_{10}, BODY: h_{11})} \\ \{ h_{10} =_q h_5 \} \end{array} \right|$

$\langle h_1, e_2\{\text{SF PROP-OR-QUES}\},$   
 $\left. \begin{array}{l} h_3: \text{"_sbiaw\_n\_coyote\_rel"(ARG0: x_4\{PNG.PER 3RD\}), \\ h_5: \text{"_copula\_v\_id\_rel"(ARG0: e_2, ARG1: x_6, ARG2: x_4), \\ h_7: \text{"exist\_q\_rel"(ARG0: x_4, RSTR: h_8, BODY: h_9), \\ h_{10}: \text{"_?u}\chi^w\_v\_go\_rel"(ARG0: e_{11}\{E.ASPECT IMPERFECTIVE\}, ARG1: x_6) \\ \{ h_8 =_q h_3 \} \end{array} \right|$

Sentences in which the general prefixes attach to a noun are also handled by this approach. The example in (110a) is analyzed as in (125). Because the predicative-noun rule creates a verb-

like element and posits an event-type variable associated with the copula relation, the irrealis prefix is able to constrain the TENSE property of that event-variable. The morphosyntactic derivation tree for this example is shown in (125a), while the corresponding semantics is presented in (125b).

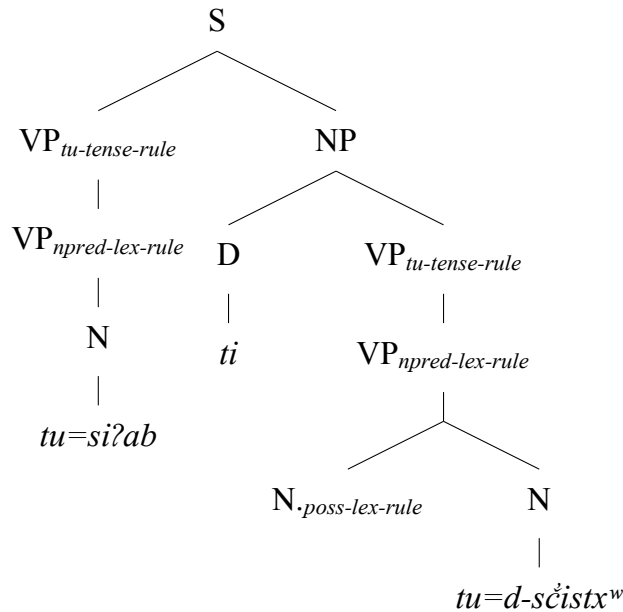
(125) a.



b.  $\langle h_1, e_2 \{ SF \text{ PROP-OR-QUES}, E. \text{ TENSE } IRR \},$   
 $h_3: \text{"_bəščəb\_n\_mink\_rel"}(\text{ARG0: } x_4 \{ \text{PNG.PER } 3RD \}),$   
 $h_5: \text{"_copula\_v\_id\_rel"}(\text{ARG0: } e_2, \text{ ARG1: } x_6, \text{ ARG2: } x_4),$   
 $h_7: \text{"exist\_q\_rel"}(\text{ARG0: } x_4, \text{ RSTR: } h_8, \text{ BODY: } h_9),$   
 $h_{10}: \text{"_ti?iʔ\_n\_dist.dma\_rel"}(\text{ARG0: } x_6),$   
 $h_{11}: \text{"exist\_q\_rel"}(\text{ARG0: } x_6, \text{ RSTR: } h_{12}, \text{ BODY: } h_{13})$   
 $\{ h_8 =_q h_3, h_{12} =_q h_{10} \} \rangle$

This attachment of general prefixes to the event variable associated with the predicative reading works in both main-predicate position as well as in argument position. The sentence in (110b), which has two past-tense markers, is analyzed as having two independent past-tense events (126b): both  $e_2$  and  $e_{16}$  have their TENSE property constrained to *past*. Effectively, this sentence is analyzed as subject-sharing between the two predications; the two copula relations share an ARG1 value ( $x_6$ ).

(126) a.



- b.  $\langle h_1, e_2 \{SF \text{ PROP-OR-QUES}, E.TENSE \text{ PAST}\},$   
 $h_3: \text{"_si?ab\_n\_honored.person\_rel"}(\text{ARG0: } x_4 \{PNG.PER \text{ 3RD}\}),$   
 $h_5: \text{"_copula\_v\_id\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_6, \text{ARG2: } x_4),$   
 $h_7: \text{"exist\_q\_rel"}(\text{ARG0: } x_4, \text{RSTR: } h_8, \text{BODY: } h_9),$   
 $h_{10}: \text{"_s'cistx^w\_n\_husband\_rel"}(\text{ARG0: } x_{11} \{PNG.PER \text{ 3RD}\}),$   
 $h_{12}: \text{"_d\_x\_1sg.poss\_rel"}(\text{ARG0: } i_{14},$   
 $\text{ARG1: } x_{13} \{PNG.PER \text{ 1ST}, PNG.NUM \text{ SG}\},$   
 $\text{ARG2: } x_{11})$   
 $h_{15}: \text{"_copula\_v\_id\_rel"}(\text{ARG0: } e_{16} \{E.TENSE \text{ PAST}\}, \text{ARG1: } x_6, \text{ARG2: } x_{11}),$   
 $h_{17}: \text{"exist\_q\_rel"}(\text{ARG0: } x_{11}, \text{RSTR: } h_{18}, \text{BODY: } h_{19})$   
 $\{ h_8 =_q h_3 \} \rangle$

System I gets certain things right. Under an analysis where nouns which bear general prefixes are also predicative, this approach captures their similarity with verbs: they expose the index of their subject to higher phrases; they bear prefixes which refer to putative properties of events. The question of whether or not these predicative uses of nouns in argument position is the right approach is one that I return to below when I compare the two systems. For the moment, I will wrap up exposition of System I by also mentioning the drawbacks which I see.

The first issue is the lack of testsuite coverage for the two sentences whose verbs combine with an object before promoting the subject. Neither of the sentences in (114) are analyzable by this approach (as written thus far), because the verbal objects aren't compatible with the NP construction. The crux of the issue is that the approach to word-order adopted in Chapter 6 attaches subjects first, then objects — a natural choice for VSO word orders. However, the implementation of this approach in binary rules added a constraint to head-complement rules whereby the subject list of the head daughter must be empty. The grammar allows for an optional subject to occur before attaching a complement, but the NP construction requires a non-empty subject in order to have an INDEX value to promote. The natural solution to this in a lexicalist framework such as HPSG is to implement a lexical rule which satisfies the subject valence constraint and exposes the index (a subject relativizing rule). However, such a rule breaks the symmetry which motivates System I: the idea was to treat nouns and verbs equally from the perspective of NP syntax. A relativizing lexical rule would effectively replace the work that the specialized NP construction is doing — there really isn't a need for both. In fact, a relativizing rule is one of the main components of System II, discussed below.

My second criticism of System I is listed above as a strength: it handles the attachment of general prefixes to nominal stems by saying that nouns may be verbal. It allows a straightforward analysis of the semantics of these prefixes as variable properties appropriate for event-type variables. While this is attractive in terms of system design, it's not completely clear that it's the right thing to do in terms of modelling the actual meaning of these prefixes. Of special note in this regard is the attachment of the past-tense marker to nouns. Langen (1996) claims that in these cases, *tu=* indicates that the noun it attaches to is deceased. This is also a topic I will have to revisit after presenting System II.

Thirdly, System I provides a highly articulated semantic representation which, while useful in some constructions, is not necessarily motivated across the board. To paraphrase the representation, every noun *X* now means something like “the one who is *X*”. This indirection appears to be semantically empty; the constraint it provides is a tautology when there is no further information added in the nouns' subject. While not necessarily harmful, such an overexpression of

variables and relations is not desirable, if it can be avoided.<sup>8</sup>

One final note before presenting an alternative system is that in order to take System I to completion, the predicative noun lexical rule really should be abandoned for a lexicon which lists base noun types as having the properties contributed by the lexical rule. I say that because the driving motivation behind this system is to capture as much syntactic similarity as possible between nouns and verbs. Forcing nouns to go through an obligatory transformation is fine as a shim to allow taking this system for a test-drive, but I think that it ought to be abandoned eventually, should System I become the preferred approach in my grammar.

### 8.3.2 *System II: Limited predication contexts*

I contrast the first system with one in which predicative context is limited. Instead of treating all nouns as underlyingly predicative, this system takes a more traditional approach in which nouns provide instance-type indices directly.

In complement to the special NP construction of System I which raises subjects out of verb-like elements, System II introduces a rule which does the same subject-relativizing work. Because this rule is intended to saturate the subject requirement of the verbs, my initial implementation of System II provides a non-branching phrase-structure rule (127).

---

<sup>8</sup>Cf. Beck (2013, §2.2.1).

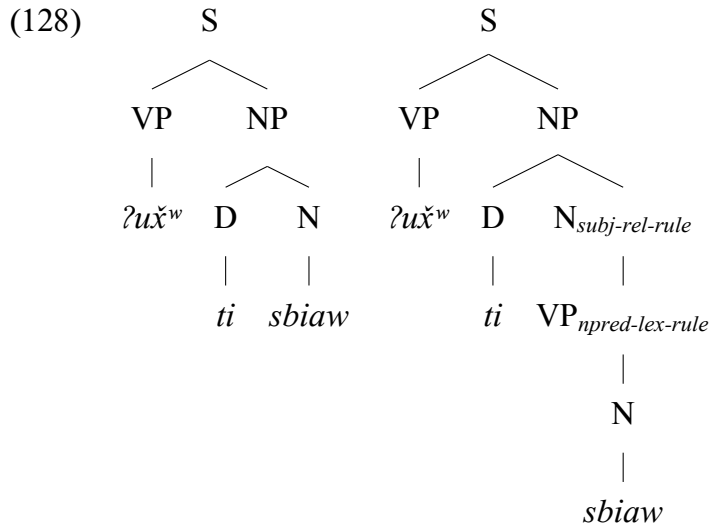
```

(127) subject-relativizing-rule := basic-unary-phrase &
 [INFLECTED #infl,
 C-CONT [RELS <! !>,
 HCONS <! !>],
 SYNSEM.LOCAL [CAT [HEAD noun & [MOD < >],
 VAL [SPR < canonical-synsem &
 [OPT -,
 LOCAL.CAT.HEAD det] >,
 SPEC < >,
 SUBJ < >,
 COMPS #comps]],
 CONT.HOOK.INDEX #x],
 ARGS < [INFLECTED #infl & infl-satisfied,
 SYNSEM.LOCAL
 [CAT [HEAD verb,
 VAL [SUBJ < [LOCAL.CONT.HOOK.INDEX #x] >,
 COMPS #comps]]] >].

```

The phrase constructed by the rule still requires a determiner, so the *SPR* feature bears an element which is *OPT -*. The rule satisfies a single *SUBJ* requirement and may apply to a verb with an unsaturated *COMPS* list. Indeed, because the output of the rule has an empty *SUBJ* list, this rule feeds the grammar's *head-complement* rule, which is desirable.

With no further constraints on the grammar, this system provides significant coverage over the testsuite. However, System II, as described so far, in fact yields a lot of ambiguity. First of all, the predicative noun rule from above applies to a noun and yields a verb with a subject requirement. Such a structure feeds the subject-relativizing rule. The result is that simple NPs now have two readings. One is the simple, intended output where the NP offers the noun's *INDEX* directly. The other is semantically identical to the analyses of System I, only with a more indirect syntax (128).

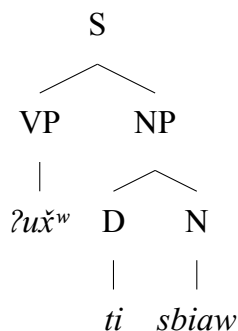


Above, I questioned the value of these vacuously relativized noun phrases when they appeared in System I. I certainly don't want them as ambiguity to a simpler representation which the grammar of System II also provides. I fixed this by adding a constraint to the predicative noun rule which keeps the HEAD value as *noun* (129). This allows them to still be predicative, but they remain nouns, which blocks them from going through the subject-relativizing rule.

(129) `n-pred-lex-rule := non-verbal-pred-lex-rule-super &`  
`[ SYNSEM.LOCAL.CAT.HEAD #head,`  
`DTR.SYNSEM.LOCAL.CAT.HEAD noun & #head ].`

After this change, predicative nouns may not feed the subject-relativizing rule, and sentences such as (107a) only have a single representation (130). For sentences such as the (a) examples of (107) and (108), the syntax (and especially the semantics) are simplified under System II as compared to System I. Instead of two events and three semantic roles to fill, there is a single event (the one introduced by the verb) and a single semantic dependency: the ARG1 of that event. The syntax and semantics of (107a) under System II are presented in (130).

(130) a.

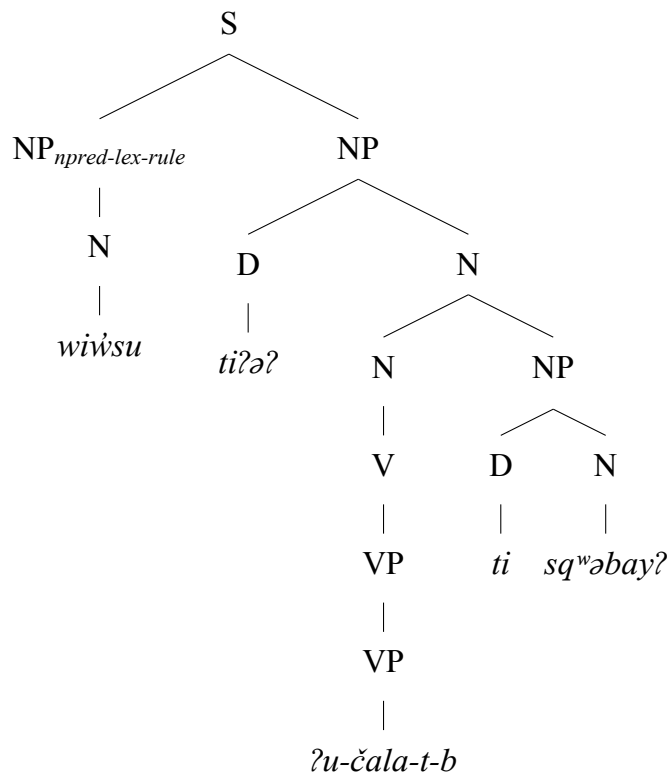
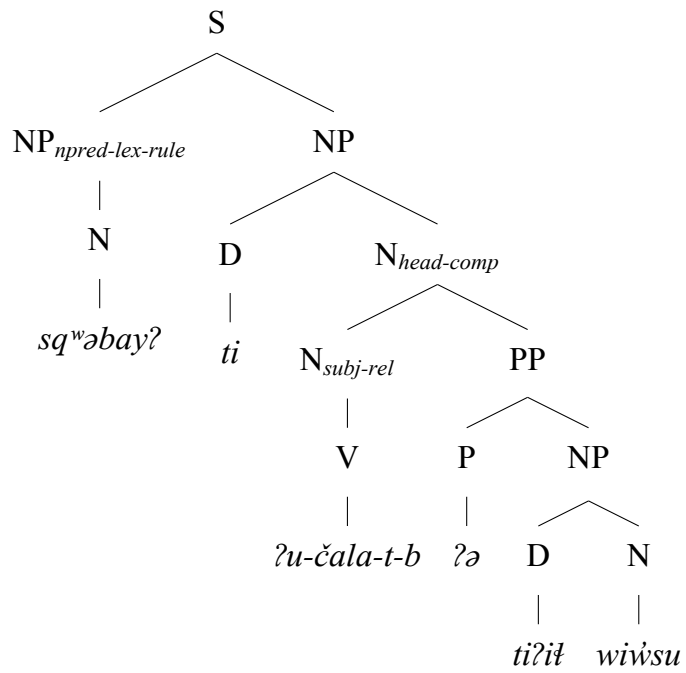


- b.  $\langle h_1, e_2\{\text{SF PROP-OR-QUES, E.ASPECT IMPERFECTIVE}\},$   
 $h_3: \text{"_?u}\chi\text{w\_v\_go\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_4\{\text{PNG.PER 3RD}\}),$   
 $h_5: \text{"_ti\_q\_unique\_rel"}(\text{ARG0: } x_4, \text{RSTR: } h_6, \text{BODY: } h_7),$   
 $h_8: \text{"_sbiaw\_n\_coyote\_rel"}(\text{ARG0: } x_4)$   
 $\{ h_6 =_q h_8 \} \rangle$

System II also handles the cases of (114), where the embedded verb has an explicit NP argument. As mentioned above, the head-complement rule is able to apply to the output of the subject-relativizing rule within the noun-phrase, so the two examples which failed to parse under System I are analyzed under System II as shown in (131) and (132). Indeed, the predicative-noun rule fires in these sentences as well, but now only on the topmost noun, which is providing the main predicate. Under System I every noun was associated with a copula relation, but now only nouns which are actually the main predicate of a sentence have them. And the semantics of these two sentences is in fact distinguished by whether or not “the children” or “the dog” provides the ARG2 of the sentential copula.<sup>9</sup>

<sup>9</sup>While it’s out of scope for this case study, it appears to me that the copula relation provided by the predicative noun rule may also be a parsimonious place to hang the focus semantics associated with such cleft-like structures. Support for information structure constraints in Matrix based grammars is provided by the information-structure analysis and library described in Song (2017).

(131)



- (132)  $\langle h_1, e_2 \{SF \text{ PROP-OR-QUES}\},$
- |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |  |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| $h_3: \text{"_sq}^w\text{əbay?_n\_dog\_rel"}(\text{ARG0: } x_4 \{PNG.PER 3RD\}),$<br>$h_5: \text{"_copula\_v\_id\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_6, \text{ARG2: } x_4),$<br>$h_7: \text{"exist\_q\_rel"}(\text{ARG0: } x_4, \text{RSTR: } h_8, \text{BODY: } h_9),$<br>$h_{10}: \text{"_ti\_q\_unique\_rel"}(\text{ARG0: } x_6, \text{RSTR: } h_{11}, \text{BODY: } h_{12}),$<br>$h_{13}: \text{"_čala\_v\_be.chased\_rel"}(\text{ARG0: } e_{14} \{E.ASPECT PERFECTIVE\},$<br><div style="text-align: center;"><math>\text{ARG1: } x_6</math>)<br/> <math>h_{13}: \text{"_t\_x\_ics\_rel"}(\text{ARG0: } e_{14}, \text{ARG1: } x_{15} \{PNG.PER 3RD\}),</math><br/> <math>h_{16}: \text{"_ti?i4\_q\_dist\_rel"}(\text{ARG0: } x_{15}, \text{RSTR: } h_{17}, \text{BODY: } h_{18}),</math><br/> <math>h_{19}: \text{"_wi}^w\text{su\_n\_children\_rel"}(\text{ARG0: } x_{15})</math><br/> <math>\{ h_8 =_q h_3, h_{11} =_q h_{20}, h_{17} =_q h_{19} \}</math> </div> |  |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
- $\langle h_1, e_2 \{SF \text{ PROP-OR-QUES}\},$
- |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |  |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| $h_3: \text{"_wi}^w\text{su\_n\_children\_rel"}(\text{ARG0: } x_4 \{PNG.PER 3RD\}),$<br>$h_5: \text{"_copula\_v\_id\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_6, \text{ARG2: } x_4),$<br>$h_7: \text{"exist\_q\_rel"}(\text{ARG0: } x_4, \text{RSTR: } h_8, \text{BODY: } h_9),$<br>$h_{10}: \text{"_ti?ə?_q\_prox\_rel"}(\text{ARG0: } x_6, \text{RSTR: } h_{11}, \text{BODY: } h_{12}),$<br>$h_{13}: \text{"_čala\_v\_be.chased\_rel"}(\text{ARG0: } e_{14} \{E.ASPECT PERFECTIVE\},$<br><div style="text-align: center;"><math>\text{ARG1: } x_{15} \{PNG.PER 3RD\}</math>)<br/> <math>h_{13}: \text{"_t\_x\_ics\_rel"}(\text{ARG0: } e_{14}, \text{ARG1: } x_6),</math><br/> <math>h_{16}: \text{"_ti?ə?_q\_prox\_rel"}(\text{ARG0: } x_{15}, \text{RSTR: } h_{17}, \text{BODY: } h_{18}),</math><br/> <math>h_{19}: \text{"_sq}^w\text{əbay?_n\_dog\_rel"}(\text{ARG0: } x_{15})</math><br/> <math>\{ h_8 =_q h_3, h_{11} =_q h_{20}, h_{17} =_q h_{19} \}</math> </div> |  |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|

Thus far, System II seems like an improvement over System I: with a simple constraint on the head value of the predication rule, this system provides a cleaner semantics for simple noun phrases while still allowing nouns to stand as the main predicate of a sentence. However, another result of the change which prevented predicative nouns from feeding the subject-relativizing rule is that predicative nouns may no longer appear in determiner phrases.

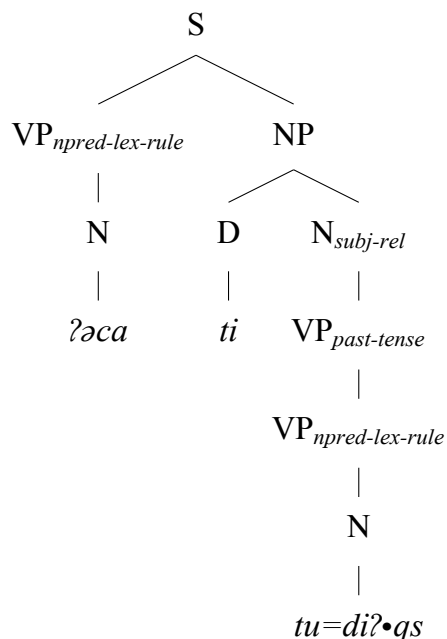
What that means in terms of testsuite coverage is that examples in which the general prefixes

attach to the noun within a determiner phrase are not analyzed. Specifically, the two sentences (110b) and (111) fail to parse. Not all examples with general prefixes on nouns fail to parse. The example sentence in (110a), however, does parse with an equivalent analysis to that given by System I: the irrealis marker constrains the event variable associated with the copula relation, specifying its TENSE property. The difference between this item and the two which fail to parse is that in this example the general prefix attaches to a predicative noun which heads the sentence.

In the examples which cannot be analyzed by this system, the past tense marker attaches to a noun which is embedded under a determiner. It's notable that the same constraint which was used to rule out what appeared to be spurious ambiguity in (107a) also rules out what would be required here to build the structures to which a past-tense marker could attach. In fact, before adding the additional head-identity constraint to the noun-predication rule, System II did parse both (110b) and (111), unambiguously. Example (133) shows the grammar's output for (111). Note the tense value specified on e16, the event associated with the copula relation built within the embedded NP. The semantic representation here is the same as the one given by System I: the two predicative noun phrases are treated as sharing a subject value. At the risk of watering-down the formal presentation, I might paraphrase this sort of reading as "[the one who X] = [the one who Y]". My use of = being due to the representation containing a single token variable, x6 in this case, for the ARG1 of both copula relations, the X and Y properties being given by the two nouns of the sentence.

In (133) (which, recall, is the output of System II *before* the head identity constraint of (129) was added), each of the two nouns goes through the predicative rule which posits a copula relation between the noun and a subject. Then, because the lower noun *tudi?qs* ("what was on the other side of the point") is able to feed the subject-relativization rule, it can expose its subject's INDEX to be picked up as the subject of the main clause. In this example, there is no ambiguity because the past-tense marker is only semantically equivalent with nouns which have undergone the predicative rule.

(133) a.

b.  $\langle h_1, e_2\{SF\ PROP-OR-QUES\},$  $h_3:"_?əca\_n\_1sg\_rel"(ARG0: x_4\{PNG.PER\ 1STPNG.NUM\ SG\}),$  $h_5:"\_copula\_v\_id\_rel"(ARG0: e_2, ARG1: x_6, ARG2: x_4),$  $h_7:"exist\_q\_rel"(ARG0: x_4, RSTR: h_8, BODY: h_9),$  $h_{10}:"\_ti\_q\_unique\_rel"(ARG0: x_6, RSTR: h_{11}, BODY: h_{12}),$  $h_{13}:"\_diqs\_n\_other.side.of.point\_rel"(ARG0: x_{14}),$  $h_{15}:"\_copula\_v\_id\_rel"(ARG0: e_{16}\{E.TENSE\ PAST\}, ARG1: x_6, ARG2: x_{14}),$  $h_{17}:"exist\_q\_rel"(ARG0: x_{14}, RSTR: h_{18}, BODY: h_{19})$  $\{ h_8 =_q h_3, h_{11} =_q h_{20}, h_{18} =_q h_{13} \}$ 

So, in System II, as written, the only way to attach a past tense marker (or any of the general prefixes which provide properties which are associated with eventuality-type variables) to a noun is to create the predicative version of that noun so that the past-tense marker has an appropriately-typed variable to constrain. However, such a construction, without further adjustment, is incompatible with a determiner because it exposes that same event-typed variable as its index; and determiners in System II want to combine with items that provide an instance-type index.

To sum up the presentation of System II, like System I, it gets certain things right. It provides a simplified semantics for nouns which are clearly in non-predicative contexts. Because it uses a phrase-structure rule to relativize verbs (instead of the determiner construction of System I), it is able to handle the relativization of verbs with complements. However, it also exposes certain drawbacks: it yields what seems to be spurious ambiguity for noun phrases which do not appear to be predicative. Such nouns are treated as both vacuously relativized and as simple. Ruling out this ambiguity, however, also rules out the application of the predicative noun rule in all embedded contexts. So examples where the predicative noun rule would have provided a reasonable attachment point for event-associated variable properties such as tense and mood are no longer able to occur in determiner phrases.

There is one further modification of System II which should be presented before I move on. The attachment of general prefixes to nouns in embedded contexts might be salvaged by reconfiguring the semantic requirements of those prefixes, either by considering non-verbal versions which contribute different meanings, or by rebuilding the variable type-hierarchy in such a way that allows properties such as tense or mood to appear on instance-type variables.

Considering the former idea, rescuing the attachment of general prefixes to nouns by having them contribute different semantics, there is precedent for the idea in the Langen's (1996) observation that when the past tense marker attaches to nouns, it indicates not that some event was in past tense, but that the individual referred to by the noun is deceased. This, in the context of System II with the prohibition on predicative nouns in embedded contexts, would provide a way to parse (111), because the lower noun could combine with the nominal version of the past tense marker instead of the verbal one.

The next question is whether such an approach is consistent with the data. I looked for examples in the corpus where past tense attaches to a noun in a context where the referent of the noun had not died, since such examples weaken the idea that a separate past-tense rule for nouns is the only way to treat nouns which bear the past-tense marker. In most of the examples I found, the relevant character in the story is unaware of the survival. For example, in Agnes James' telling of *Basket Ogress*, the children who are presumed to have been killed by the titular

character are referred to in past tense by those who presume them dead just before they arrive (134).

- (134) *g<sup>w</sup>əl ʔuχ<sup>w</sup>tub ti bə<sup>k</sup> tubədbədaʔləp*  
*g<sup>w</sup>əl ʔuχ<sup>w</sup>-tx<sup>w</sup>-b ti bək<sup>w</sup> tu=bəd-bədaʔ-ləp*  
 SCONJ go-ECS-PASS DET all PST=DSTR-offspring-2PL.POSS

‘ “and she has taken all of your children.” ’ (reported speech of Little Hunchback)

Basket Ogress (AJ)

As the children reappear, the past-tense marker is removed (135). Notably, however, Beck’s gloss here is “former”. In fact, in the story, these children have moved away from their village. What I’m getting at is that there really may be two meanings here which are both compatible with the facts of the story: in meaning (1), the past tense marker means dead, and the usage provides a kind of rhetorical flourish where the people see the “dead children”, only to realize that they are not in fact dead; in meaning (2), the past tense marker does not actually mean dead, but indicates past tense on some event-like property, a reading which seems to correspond with the gloss “former children of the village.”

- (135) a. *huy g<sup>w</sup>əl šuucəb tiʔəʔ dəx<sup>w</sup>əstəlil ʔə tiʔəʔ tuwiwsu*  
*huy g<sup>w</sup>əl šuʔ-c-b tiʔəʔ dəx<sup>w</sup>=ʔas-ʔəʔlil ʔə tiʔəʔ tu=wiwsu*  
 SCONJ SCONJ see-ALTV-PASS DET ADNM=STAT-live PR DET PST=children

‘And then they see the former children of the village.’

- b. *ti<sup>l</sup>əb k<sup>w</sup>i sʔəʔ ʔə ti wiwsu*  
*ti<sup>l</sup>əb k<sup>w</sup>i s=ʔəʔ ʔə ti wiwsu*  
 immediately DET NMLZ=COME PR DET children

‘the coming of the children is sudden.’

Basket Ogress (AJ)

Complicating matters for this sort of corpus study, however, is that in Lushootseed, the usage of past and future tense markers are completely optional at the sentential level but are used instead at the discourse level to indicate point-of-reference time as narratives progress from episode to episode (Bates and Hess, 2001; Bates, 2002). Nevertheless, the question at hand for me is whether my Lushootseed grammar might be augmented by a separate set of general prefixes which are semantically tailored for nouns. Assuming, for the moment, that Lushootseed requires at least a second form of the past-tense prefix which does not constrain the tense property of an event-type variable but instead indicates that the referent is deceased, this still fails to completely address the data in the testsuite.

System II, augmented only by a noun-only version of the past-tense marker, can now analyze the example in (110b) (copied here for convenience as (136)). The noun *tudščistx<sup>w</sup>* can now only mean “deceased husband” since the predicative noun rule is still prevented from feeding the subject relativizing rule. The higher noun must still pass through the predicative noun rule in order to take a subject, and since this rule (under System II) preserves the head value, only the nominal past-tense marker can attach to the noun yielding a semantics which might be glossed “deceased person-of-rank”.

- (136) *tusi?ab*                      *ti tudščistx<sup>w</sup>*  
 tu=si?ab                      ti tu=d-ščistx<sup>w</sup>  
 PST=person.of.rank    DET    PST=1SG.POSS-husband

‘My former husband was a man of rank.’

LR1 p.81

That seems not unreasonable to me. However, the new parse for (111) (copied here as (137)) bears an apparently nonsensical reading. As in (136) only the noun-attaching past-tense marker with the meaning “deceased” can attach within the noun phrase, so the phrase *ti tudi?qs* is assigned a meaning which might be glossed “deceased other side of the point”. This is certainly not correct. One caveat which applies to this example is that Lushootseed grammar does provide a special syntactic position for location-phrases. So, it may be possible to rescue this particular

example by saying that *diʔqs* is not a typical noun, but a something more like an adjective or adverb.

- (137) ʔəca ti tu=diʔqs  
 ʔəca ti tu=diʔ•qs  
 1SG DET PST=other.side•point

‘I am the one who was on the other side of the point.’

Special treatment for location phrases may suffice for this example, so under the assumption that the details of location phrases can be sorted out in such a way as to keep the rest of System II (as presented so far) unchanged, there is now an awkward asymmetry on two dimensions which is difficult to motivate. On one dimension, some authors have argued for a separate semantics associated with *tu=* when it attaches to nouns. This would seem to rescue a hypothesis that predicative nouns should not feed the subject relativization rule, because it provides a way to parse embedded nouns which bear the past tense prefix without sending them through the predicative noun rule. The asymmetry, however, is that there is no such alternative semantics for the other general prefixes. The testsuite contains example (110a), copied here for convenience as (138), where the irrealis (or future) marker attaches to a noun in predicate position. Unless I posit a new, noun-attaching form for the irrealis marker, this example fails to parse with the modifications to System II which are under discussion. If I do posit a new noun-attaching version of the irrealis, what would its semantics be? The same situation must eventually be addressed for the other general prefixes (analyzed in my current system as constraining mood properties of event-type variables). Even though these prefixes don’t appear in the testsuite, they do appear in the corpus on nouns even in embedded contexts (139).<sup>10</sup>

---

<sup>10</sup>Recall that possessive marking is only available on nouns, so *ʔuqədʔaxs* is clearly nominal.

(138) *ʎubəščəb tiʎiʎ*

ʎu=bəščəb tiʎiʎ

IRR=mink DIST.DEM

‘That one will become a mink.’

(139) *put haʎʎ ti ʎuqədʎaʎs*

put haʎʎ ti ʎu=qədʎaʎ-s

really good DET HAB=intestines-3po

‘Her intestines are very nice.’

Mink and Tutyika: Second telling (ML)

The other asymmetry which exists in System II regards the contexts in which nouns may be predicative. System II prevented multiple analyses of simple noun phrases which appear with determiners by preventing them from feeding the subject relativization rule. However, examples such as (137) suggest to me that embedded nouns can be predicative, necessitating relativization. Given the fact that (137) concerns a location phrase, I looked for other examples in the corpus where past tense appears on nouns which are within noun phrases, but which didn’t concern locations. Additionally, I restricted my search to examples where it was implausible that the past tense marker meant “dead”. The first set of examples which I turned up in this search were found in Martha Lamont’s telling of Seal Hunters. In one episode of the story, the brothers who have been carried off to the far away land notice large amounts of dentalia shells piled in front of the dwarves’ houses. The dentalia are referred to without any past tense marking the first 5 mentions (140). On lines 434 – 436 and 438, the dentalia shells are introduced (and on line 439, the fact that the brothers are desirous of the shells). After this several more events take place before the dentalia are returned to. On line 570, the brothers gather up the shells.

- (140) a. *huy put ?aspuk<sup>w</sup>puk<sup>w</sup>əb ti?ə? s?uləḥ čəg<sup>w</sup>adi?*  
 huy put ?as-puk<sup>w</sup>-puk<sup>w</sup>əb ti?ə? s?uləḥ čəg<sup>w</sup>•adi?  
 SCONJ really STAT-DSTR-piled.up DET dentalia seaward•side

‘Then, a lot of dentalia are piled up in front of the houses.’

- b. *ti?ə? λus?uləḥ diḥ stabig<sup>w</sup>s*  
 ti?ə? λu=s?uləḥ diḥ stabig<sup>w</sup>s  
 DET HAB=dentalia FOC valuables

‘These dentalia are valuable goods.’

- c. *?al k<sup>w</sup>ədi? tuha?k<sup>w</sup> ti?iḥ s?uləḥ*  
 ?al k<sup>w</sup>ədi? tu=ha?k<sup>w</sup> ti?iḥ s?uləḥ  
 LOC DET PST=long.time DET dentalia

‘They had been there a long time.’

- d. *tuḥ<sup>w</sup> (h)uy hik<sup>w</sup>ik<sup>w</sup> (h)aḥ s?uləḥ*  
 tuḥ<sup>w</sup> huy hik<sup>w</sup>-hik<sup>w</sup> haḥ s?uləḥ  
 just SCONJ DSTR-big good dentalia

‘There were lots of dentalia.’

- e. *hay g<sup>w</sup>əl ?a q<sup>w</sup>u?ədəx<sup>w</sup> əlg<sup>w</sup>ə? ti?ə? s?uləḥ*  
 hay g<sup>w</sup>əl ?a q<sup>w</sup>u?-t=ax<sup>w</sup> əlg<sup>w</sup>ə? ti?ə? s?uləḥ  
 SCONJ SCONJ INTJ gathered-ICS=now PL DET dentalia

‘And so they gather the dentalia shells.’

Then, on line 571, the noun *s?uləḥ* (dentalia shells) appears with past tense (141a). The shells are referred to nine more times without a tense marker as they are packed up and loaded for a trip. Then, on line 663, they appear again in past-tense (141b). The word for dentalia is etymologically related to the verb *?uləḥ* which means “be gathered”. I have preserved Beck’s

glossing in presenting this example in which he has represented this etymological substructure explicitly in the segmented lines.<sup>11</sup> After this, the dentalia shells are mentioned seven more times in the story. The next two times, they are marked with past tense. After that, in reference to the dentalia that sometimes wash ashore (to this day) near where the spill occurred, the present tense is used.

- (141) a. *g<sup>w</sup>əl q<sup>w</sup>u?ədax<sup>w</sup> əlg<sup>w</sup>ə? tə diʔax<sup>w</sup> haʔt ʔə k<sup>w</sup>i tusʔuləḥ*  
*g<sup>w</sup>əl q<sup>w</sup>u?-t=ax<sup>w</sup> əlg<sup>w</sup>ə? tə diʔ=ax<sup>w</sup> haʔt ʔə k<sup>w</sup>i tu=sʔuləḥ*  
 SCONJ gathered-ICS=NOW PL DET FOC=NOW good PR DET PST=dentalia

‘And they gather what were good dentalia shells.’

- b. *g<sup>w</sup>əl huy k<sup>w</sup>ʔax<sup>w</sup> tiʔiʔ tusʔuləḥ<sup>w</sup>s əlg<sup>w</sup>ə?*  
*g<sup>w</sup>əl huy k<sup>w</sup>ʔ=ax<sup>w</sup> tiʔiʔ tu=s=ʔuləḥ<sup>w</sup>=s əlg<sup>w</sup>ə?*  
 SCONJ SCONJ spill=NOW DET PST=NMLZ=gathered=3.POSS PL

‘And then their dentalia spilled.’

- c. *...ʔə tiʔiʔ tusʔuləḥ<sup>s</sup> əlg<sup>w</sup>ə?*  
*...ʔə tiʔiʔ tu=sʔuləḥ<sup>s</sup>-s əlg<sup>w</sup>ə?*  
 ...PR DET PST=dentalia-3.POSS PL

‘There to seaward from where they had been, their dentalia spilled.’

To me, these examples indicate that past tense must be able to attach on nouns without meaning “deceased” even for nouns which do not represent locations. Analytically, this reading can be captured in a grammar by allowing the predicative noun rule to fire, even in embedded contexts, and then attaching the past-tense marker to the event-type variable associated with the copula relation. This is the relative-clause-like reading that is allowed by System I but not System II. Indeed, Beck’s free translation of (141a) captures the corresponding English paraphrase “what

<sup>11</sup>I suspect that this representation may have been chosen in order to differentiate this past tense marking on a noun as distinct from the examples where it means “deceased”.

were good dentalia shells.”

There are other examples in the corpus where past tense on nouns cannot refer to being dead. In Alice William’s telling of Basket Ogress, a woman and child escape from the ogres. Upon discovering their escape, they refer to the people (their food) in past tense (142). The people are not dead (they would have been killed had they not escaped). On the other hand, their state as prospective food is past. This, to me, fits nicely with a treatment where the past tense marker is attached to the noun after the application of the predicative lexical rule which introduces the event-type variable.

- (142) *g<sup>w</sup>əl ʔuɫi<sup>w</sup> tiʔi<sup>l</sup>ə tusʔətəd čət*  
*g<sup>w</sup>əl ʔu-ɫi<sup>w</sup> tiʔi<sup>l</sup>ə tu=sʔətəd čət*  
 SCONJ PFV-escape DET PST=food 1PL.POSS

‘And our food has escaped.’

Basket Ogress (AW)

There are a few other such examples in the corpus, but I think that what’s left to decide, from the perspective of grammar implementation, is whether the “dead” reading deserves its own semantics. On that topic, it is less plausible that I can argue from examples, but if it’s plausible to chalk up the deceased readings to a pragmatic interpretation of “the one who was X”, then the grammar can move forward without introducing further ambiguity. The ambiguity I’m talking about is that if I must allow the predicative noun rule to attach to embedded nouns and to support application of the past prefix (examples such as (141) and (142) make me think that I must), then allowing these forms to also be analyzed with low attachment of the past prefix (to the noun, yielding the “dead” reading), makes them ambiguous. In a grammar with both attachments, forms such as *tudsčistx<sup>w</sup>* (136) are ambiguous between the deceased reading and the former reading — the former reading being the one with low attachment of the past prefix to a noun (the “dead” semantics would be contributed via an elementary predication), the latter being high attachment of the past prefix after applying the predicative rule with semantics effected as a constraint on the tense property of the event variable.

To me, the data suggests that System II is incorrect limiting nominal predication contexts to just the main predicative position. Doing so was an attempt to constrain what appears in most sentences to be spurious ambiguity: simple noun phrases in argument position have two readings otherwise: the streamlined one in which the noun simply contributes its index to the determiner, and the relativized one where the noun becomes predicative and contributes its subject's index via the subject-relativization rule. The ambiguity appears spurious because without further information or constraint on the subject of the noun, the index of the subject of the noun is a kind of useless abstraction over the index of the noun itself: “ $y$  where  $isa(e, y, x)$ ” is no more informative than just plain “ $x$ ” given no further information about  $e$  or  $y$ . Nevertheless, examples with past tense marking on the noun phrase in these sorts of contexts provide a motivation for the abstraction: “ $y$  where  $isa(e, y, x) \ \& \ past(e)$ ” is more informative than just “ $x$ ”.<sup>12</sup>

The next section provides a summary comparison of the two systems. This is followed by a summary of the results against the testsuite.

#### **8.4 Discussion and Results on the Testsuite**

The two systems for handling non-verbal predicates which I presented in the last section each have some strengths and some weaknesses. System I, where every content word provides a predicate and a corresponding event-like variable provides parses for all of the tricky cases which vex System II — specifically, the cases where nouns appear to be predicative in embedded contexts. System I does this without creating any spurious ambiguity. However, it does so by positing an abstraction in the semantics which is often vacuous. With some modifications, System II can accommodate all of the data which is handled by System I, but these modifications cause the grammar to yield extra parses which map to the same semantics.

Overall, I consider this scenario to illustrate a win for the metagrammar approach of the CLIMB methodology. Because I'm not completely satisfied with either option, I can choose to retain them both for the time being, and this is what I have done. In the next chapter, I describe

---

<sup>12</sup>Bender (pc) points out that another potential approach would seek to rule out the predicate-nominals unless they bear general prefixes or appear in sentence-initial (predicate) position.

| <b>'lushootseednomdets/2019-06-29_23:12:05_utc/nominal-predicates/19-06-29/lkb' Coverage Profile</b> |                    |                       |                    |                      |                          |                      |                         |
|------------------------------------------------------------------------------------------------------|--------------------|-----------------------|--------------------|----------------------|--------------------------|----------------------|-------------------------|
| <b>Aggregate</b>                                                                                     | <b>total items</b> | <b>positive items</b> | <b>word string</b> | <b>lexical items</b> | <b>distinct analyses</b> | <b>total results</b> | <b>overall coverage</b> |
|                                                                                                      | #                  | #                     | $\phi$             | $\phi$               | $\phi$                   | #                    | %                       |
| $5 \leq i\text{-length} < 10$                                                                        | 2                  | 2                     | 5.50               | 7.50                 | 0.00                     | 0                    | 0.0                     |
| $0 \leq i\text{-length} < 5$                                                                         | 15                 | 15                    | 2.53               | 2.71                 | 1.00                     | 14                   | 93.3                    |
| <b>Total</b>                                                                                         | <b>17</b>          | <b>17</b>             | <b>2.88</b>        | <b>3.31</b>          | <b>1.00</b>              | <b>14</b>            | <b>82.4</b>             |

(generated by [incr tsdb()] at 30-jul-2019 (21:52 h))

Table 8.2: The Testsuite Results for System I

| <b>'lushootseedmrs/2019-06-30_15:56:11_utc/nominal-predicates/19-06-30/lkb' Coverage Profile</b> |                    |                       |                    |                      |                          |                      |                         |
|--------------------------------------------------------------------------------------------------|--------------------|-----------------------|--------------------|----------------------|--------------------------|----------------------|-------------------------|
| <b>Aggregate</b>                                                                                 | <b>total items</b> | <b>positive items</b> | <b>word string</b> | <b>lexical items</b> | <b>distinct analyses</b> | <b>total results</b> | <b>overall coverage</b> |
|                                                                                                  | #                  | #                     | $\phi$             | $\phi$               | $\phi$                   | #                    | %                       |
| $5 \leq i\text{-length} < 10$                                                                    | 2                  | 2                     | 5.50               | 7.50                 | 1.00                     | 2                    | 100.0                   |
| $0 \leq i\text{-length} < 5$                                                                     | 15                 | 15                    | 2.53               | 2.80                 | 1.00                     | 13                   | 86.7                    |
| <b>Total</b>                                                                                     | <b>17</b>          | <b>17</b>             | <b>2.88</b>        | <b>3.35</b>          | <b>1.00</b>              | <b>15</b>            | <b>88.2</b>             |

(generated by [incr tsdb()] at 30-jul-2019 (21:54 h))

Table 8.3: The Testsuite Results for System II

a system evaluation in which the grammar system was run against the four texts at the back of LR1. Since I retain both these systems for non-verbal predicates, I was able to customize a grammar for each and run the experiment twice. This allowed me to further compare the two systems. The results of that exercise are detailed in that chapter.

What remains for this section is to present the results of the two systems against the testsuite. The results for System I are presented in Table 8.2. Then, System II results are found in Table 8.3.

## Chapter 9

### EVALUATION

This chapter presents an evaluation of the metagrammar system which takes the form of a coverage report and error analysis on the four stories by Edward Sam which are included at the end of Lushootseed Reader Volume 1 (Hess, 1995, LR1). These stories formed a “development” corpus and are not-held out data. In that way, this evaluation chapter presents an effort which is not a true experimental evaluation but instead simply a natural step in the continuing development of the system — a further exercise in grammar engineering.

The case study of Chapter 8 left off by comparing the strengths and weaknesses of two implemented systems which are intended to capture both non-verbal predicates as well as verb-phrases which are embedded in determiner phrases. This evaluation exercise continues that comparison by exposing both systems to a larger set of data than just that which was included in the targeted testsuite constructed for that case-study.

In addition to the development work described in previous chapters, there were many other, smaller implementations of grammatical phenomena which I built into the metagrammar system over the course of development. Some of these implementations were specifically targeted at phenomena I observed in this development corpus, so I take the opportunity in this chapter to present that work as well. Overall, this evaluation exercise provides a way to benchmark the current system’s functionality. The error analysis here yields a “what’s next” list for future development.

The structure of this chapter is as follows: I begin with the other notable implementations which I built in preparation for this evaluation. These implementations were driven by the goal of providing maximal plausible coverage with minimal effort — I was hunting for low-hanging fruit. After that, I present a coverage and ambiguity report for both of the non-verbal predicate

systems of Chapter 8 for each of the four stories. This section compares the systems' output (where they differed), looking at particular examples that one system or the other excelled or failed on. The final content section provides a detailed error analysis. There are two subsections: one dedicated to failures to parse tokens lexically, the other to general parse failures, given lexical coverage. After this, I summarize the content of this chapter and conclude.

### 9.1 *Evaluation Context: other notable implementations*

Beyond the implementations of grammatical phenomena which have been described in previous chapters (Chapter 6, Chapter 7, Chapter 8), I did some further development work on the grammar in support of gaining coverage over natural language text. This section provides an overview of that work.

#### 9.1.1 *Sentential Conjunctions*

The naturally flowing stories of the Lushootseed corpus I have access to contain many sentences which begin with sentential connectives which are often glossed “and then” or “then”. These were mentioned in Chapter 6 as an aspect of word order which would need to be accounted for in order to properly place the sentence-second clitics in real sentences. One of the examples from Chapter 6 is repeated below as 143.

- (143) *g<sup>w</sup>əl tuχ<sup>w</sup> ?ashudtubax<sup>w</sup>*  
*g<sup>w</sup>əl tuχ<sup>w</sup> ?as-hud-tx<sup>w</sup>-b=ax<sup>w</sup>?*  
 SCONJ just STAT-burn-ECS-PASS=NOW

‘They just build a fire.’

Basket Ogress (ML)

My approach was essentially the one I described in that chapter. I added a new class of adverbs into the choices file which target *S*-type constituents. Then, programmatically, when the grammar is customized (i.e. built from choices file input), lexical entries are assigned one of two supertypes, either the one for pre-verbal scopal adverbs (such as *tuχ<sup>w</sup>* (“just”) in (143)) or

the one for sentential conjunctions (such as *gʷəl* (“and then”) in (143)) based on their listing in the choices file. Because most of the supertype information for the adverbs has been extracted into a common supertype, the discriminant which handles their word-order difference is simply whether or not their MOD value is LIGHT (see Chapter 6 for details).

### *NP Coordination*

The Grammar Matrix provides a coordination library (Drellishak and Bender, 2005) which offers out-of-the box functionality for setting up various coordination strategies within a grammar (Drellishak and Bender, 2005). Grammars may implement multiple coordination strategies and assign them to particular constituent types. Among other coordination phenomena, Lushootseed exhibits NP coordination around an explicit conjunction *?i*.

The entry for *?i* in the *Lushootseed Dictionary* (Bates et al., 1994) lists examples which seem to exhibit more than one syntactic pattern. In (145a), a polysyndeton<sup>1</sup> strategy joins full NP coordinands. In (145b), however, a proper name is conjoined with some constituent in what will almost certainly need to be analyzed as some type of ellipsis. The grammar of (145b) falls outside of the scope of the phenomena that can be treated by the customization system out-of-the-box (since there is no library for describing ellipsis strategies). In (145c), the pattern looks like omnisyndeton: each of the coordinands is marked with a conjunction. However, (145c) is also notable because the order of the determiner and conjunction is not consistent between the two coordinands. Like (145b), a full treatment of the complexities exhibited by (145c) lies outside

---

<sup>1</sup>Drellishak and Bender (2005) contrast the terms omni-, poly- and mono-syndeton in describing the co-occurrence patterns of conjunctions with coordinands. An omnisyndeton pattern is one in which each coordinand is associated with an explicit conjunction (144a). Polysyndeton marks all but the first coordinand (i.e. the conjunctions appear between each pair of coordinands) as in (144b). In a monosyndeton pattern, some particular coordinand is distinguished by its bearing of the explicit conjunction (144c).

- (144) a. and A and B and C  
 b. A and B and C  
 c. A B and C

of the scope which I have allowed for myself in these rudimentary implementations. Then, in (145d), yet another pattern is attested. In this latter case, to be fair, the vowel lengthening on the conjunction (and the ellipses in the gloss) indicates that the speaker may have been connecting this entire line with the subsequent one.

- (145) a. *tutčisəb čəł ʔə tiʔəʔ stubš ʔi tsiʔiʔ čəgʷas ʔi tiʔiʔ*  
 tu=ʔčil-s-b čəł ʔə tiʔəʔ stubš ʔi tsiʔiʔ čəgʷas ʔi tiʔiʔ  
 PST=arrive-ICS-PASS 1PL.SBJ PR DET man CONJ DET.F wife CONJ DET  
*bədbədaʔs*  
 CVC-bədaʔ-s  
 DSTR-offspring-3POSS

‘A man and (his) wife and their children came to us.’

- b. *tušudxʷ čəł ʔi mali*  
 tu=šut-dxʷ čəł ʔi mali  
 PST-see-DC 1SG.SBJ CONJ mary

‘Mary and I saw it.’

- c. *ʔaʔ tiʔəʔ ʔi təqxʷ ʔi tiʔəʔ suqʷaʔs*  
 ʔaʔ tiʔəʔ ʔi təqxʷ ʔi tiʔəʔ suqʷaʔ-s  
 be.there DET CONJ beaver CONJ DET younger.sibling-3.POSS

‘There was beaver and his younger brother.’

[14]bat:hes:hil:1994

- d. *sʔətəds sʔətəds sʔətəds ʔiʷ*  
 s=ʔətəd=s s=ʔətəd=s s=ʔətəd=s ʔi  
 NMLZ=food=3.POSS NMLZ=food=3.POSS NMLZ=food=3.POSS CONJ

‘And eat and eat and eat, and ...’

Bear and Fishhawk (ES)

Ruling out the patterns of (145b) – (145d) due to their complexity, leaves me with the

polysyndeton construction of (145a), which I implemented by using the coordination library to enable a coordination strategy for NPs, using the form *?i* as the conjunction. This implementation allows the grammar to parse some of the conjoined NPs found in the evaluation materials, such as (146).

- (146) *huy tük<sup>w</sup>tubəx<sup>w</sup> ti?i? bibšćəb ?i ti?i? su?suq<sup>w</sup>a?s*  
*huy tük<sup>w</sup>-tx<sup>w</sup>-b=ax<sup>w</sup> ti?i? bi-bəšćəb ?i ti?i? su?-suq<sup>w</sup>a?-s*  
 SCONJ go.home-ECS-PASS=NOW DET ATTN-mink CONJ DET ATTN-younger.sibling-3.POSS  
 ‘Then Little Mink and his little younger cousin are taken home.’ Mink (ES)

### *Attributive adjectives*

In addition to the predicative adjectives (mentioned briefly in Chapter 8), Lushootseed grammar also has attributive adjectives which may appear prenominally (147). The customization system has support for attributive adjectives, so I was able to enable these simply by listing them in the lexicon and marking them as such. These lexical entries are integrated with the non-verbal-predication handling described in Chapter 8, so any such listed attributive adjective is also able to appear as a main predicate.

- (147) *g<sup>w</sup>əl ?uχ<sup>w</sup>tub ti bə<sup>k</sup> tubədbəda?ləp*  
*g<sup>w</sup>əl ?uχ<sup>w</sup>-tx<sup>w</sup>-b ti bə<sup>k</sup> tu=bəd-bəda?-ləp*  
 SCONJ go-ECS-PASS DET all PST=DSTR-offspring-2PL.POSS  
 ‘and she has taken all of your children.’ Basket Ogress (AJ)

However, as with NP coordination discussed above, the out-of-the-box strategy for attributive adjectives doesn’t completely cover the syntactic possibilities. In (148a), the adjective appears outside of the determiner. In (148b), the adjective appears postnominally. Neither of these constructions is covered in my rudimentary implementation.

- (148) a. *huy g<sup>w</sup>əl ?atəbədəx<sup>w</sup> bə·k<sup>w</sup> ti?ə? tubədbədə?s əlg<sup>w</sup>ə?*  
*huy g<sup>w</sup>əl ?atəbəd=ax<sup>w</sup> bək<sup>w</sup> ti?ə? tu=CVC-bəda?-s əlg<sup>w</sup>ə?*  
 SCONJ SCONJ die=now all DET PST=DSTR-offspring-3.POSS PL

‘and then all her children died, too.’

Basket Ogress (AJ)

- b. *tuha?tdali?əd čəx<sup>w</sup> ?ə k<sup>w</sup>i ?aci?talbix<sup>w</sup> čit dəx<sup>w</sup>ə?əλ[s]*  
*tu=ha?ł-t•ali?əd čəx<sup>w</sup> ?ə k<sup>w</sup>i ?aci?talbix<sup>w</sup> čit dəx<sup>w</sup>=lə=?əλ’s*  
 IRR=good-ICS•food 2SG.SBJ PR DET people near ADN=PROG=come=3.POSS

‘You will be good food for the people nearby who are coming.’

Mink and Tutyika (ML)

### *Reduplication*

Lushootseed has more than one reduplication strategy. Reduplication is one of the grammatical phenomena which may apply to nouns, verbs, or other lexical categories (see §2.4.2). Like the general prefixes which also apply to both nouns, verbs and others, the integration of the semantics indicated by these reduplication patterns is not straightforwardly mappable to the variable properties which are provided by default in Grammar Matrix grammars. For example, CVC reduplication is usually glossed by Beck as marking a distributive property. When a noun is marked as distributive, it often indicates plurality, but on a verb it may indicate an iterated action or an action that is distributed across spatial locations (Beck, n.d., 320). A full study of these patterns and their semantics is more than I can accommodate here, but I was able to add rudimentary support for distributive, attentuative and diminished effectiveness reduplication patterns by creating morphological markers that stand in for them, and then writing some code in the metagrammar to associate these morphemes with a contribution of representative elementary predications. Like the analysis I argued for in Chapter (7) for valence-increasing affixes, these elementary predications share their ARG0 and LBL values with the verb they attach to. The idea behind this is to represent what semanticists call predicate-modification. The image in (149)

shows the semantics associated with *cutcut* (“say again and again”) after implementing this approach to reduplication.

$$(149) \quad \langle h_1, e_2, \left. \begin{array}{l} h_3: \text{"\_cut\_v\_speak\_rel"}(\text{ARG0: } e_2, \text{ ARG1: } x_4), \\ h_3: \text{"\_cvc\_x\_dstr\_rel"}(\text{ARG0: } e_2) \\ \{ \} \end{array} \right| \rangle$$

The approach of using elementary predications is, in principle, more easily extensible to nouns than one which uses variable properties. All that really needs to happen is that the type constraint on the ARG0 value of the grammatical EP needs to be the supertype of *event* and *instance* (and this is already the case, in my implementation). However, due to the low-attachment point of reduplication (nearest to the root), I wasn’t easily able to integrate these forms into the morphotactic pipeline without a major refactoring.<sup>2</sup> For this reason, I opted to only provide this rudimentary reduplication analysis to verbs. Nouns with reduplication marking are unanalyzable in my current system.

### *Possessive Marking*

As mentioned in Chapter 8, Lushootseed marks possession on nouns with an asymmetric paradigm. First person singular and second person singular are marked with prefixes (*d-* and *ad-*, respectively); second person plural and third person (of underspecified number) are marked with suffixes (*-ləp* and *-s*, respectively) but first person plural is marked with a stand-alone word which is homophonous with the corresponding subject clitic (*čəł*). I provide an implementation for the affixal forms of this paradigm using an approach which is similar to the one I presented above for reduplication. I used the morphotactic library (Goodman, 2013) to define the position classes and the morphological rules and their morphotactic requirements and then added

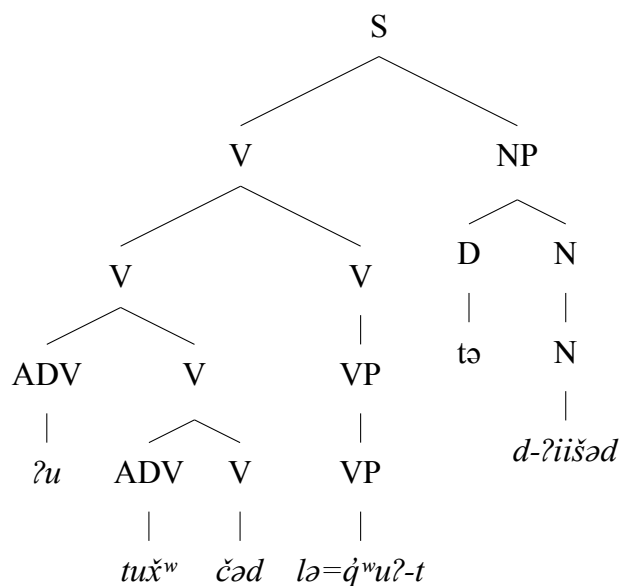
---

<sup>2</sup>It’s not entirely clear to me at the time of this writing whether the Grammar Matrix morphotactic library would support a pipeline which allows a position class to apply to either noun or verb while allowing the output of any rules in this position class to tee out into separate pipelines which must include only nouns or only verbs. My intuition is that it would, that I simply need to reconfigure my current approach to morphotactics.

semantics to the rules in my metagrammar.

The semantics of possessive marking is treated as a two-place relation: ARG1 corresponding to a possessor, ARG2 to the possessed. Thus, the index of the noun itself provides the value of ARG2, and the PNG properties of ARG1 are constrained as appropriate for the morpheme in question. The syntactically positioned marker for first person plural was left out-of-scope for this evaluation. The material in (150) shows the grammar’s syntax and semantics for an example from “Coyote and Big Rock” (ES) which exhibits first person singular possessive marking on a noun.

(150) a.



- b.  $\langle h_1, e_2 \{E.ASPECT \textit{PROGRESSIVE}\},$   
 $h_3: \textit{"_?u_sconj\_interjection\_rel"}(ARG0: e_5, ARG1: h_4),$   
 $h_6: \textit{"_tu\check{x}^w\_r\_merely\_rel"}(ARG0: e_8, ARG1: h_7),$   
 $h_9: \textit{"_q^wu?v\_gather\_rel"}(ARG0: e_2, ARG1: x_{10} \{PNG.PER 3RD\}),$   
 $h_9: \textit{"_t\_x\_ics\_rel"}(ARG0: e_2, ARG1: x_{11} \{PNG.PER 1ST, PNG.NUM SG\}),$   
 $h_{12}: \textit{"_t\check{a}\_q\_neutral\_rel"}(ARG0: x_{10}, RSTR: h_{13}, BODY: h_{14}),$   
 $h_{15}: \textit{"_?ii\check{s}\check{a}d\_n\_relatives\_rel"}(ARG0: x_{10}),$   
 $h_{16}: \textit{"_d\_x\_1sg.poss\_rel"}(ARG0: i_{18}$   
 $ARG1: x_{17} \{PNG.PER 1ST, PNG.NUM SG\}$   
 $ARG2: x_{10})$   
 $\{ h_4 =_q h_{19}, h_7 =_q h_{19}, h_{13} =_q h_{15} \} \rangle$

### *Now enclitic*

Another morphological element which I treated as a contentful suffix is the temporal enclitic =*ax<sup>w</sup>*. This marker is used within a discourse to mark a change of state or condition (Bates, 1999) and is often glossed “now” in English translations.<sup>3</sup> As with my treatments of reduplication on verbs and possessive marking on nouns, I used the morphotactics system to define a position class and then used the metagrammar to attach semantics to the rule. In contrast to the reduplication implementation, the outermost attachment point of the temporal enclitic position class facilitated my integration of this marker with both noun and verb morphology pipelines. So my grammar allows either an event-type or an instance-type ARG0 value for the *\_ax<sup>w</sup>\_x\_now\_rel* predicate that the rule contributes. The examples in (151) show the MRS associated with verbal attachment (151a) and nominal (151b) attachment.

- (151) a.  $\langle h_1, e_2,$   
 $h_3: \textit{"_?u\check{x}^w\_v\_go\_rel"}(ARG0: e_2, ARG1: x_4),$   
 $h_3: \textit{"_ax^w\_x\_now\_rel"}(ARG0: e_2)$   
 $\{ \} \rangle$

<sup>3</sup>See also Chapter 6 of Zahir (2018).

- b.  $\langle h_1, e_2,$   
 $\left. \begin{array}{l} h_3: \text{"\_b\check{a}\check{s}\check{c}\check{a}b\_n\_mink\_rel"}(\text{ARG0: } x_4), \\ h_3: \text{"\_ax}^w\text{\_x\_now\_rel"}(\text{ARG0: } x_4), \\ h_5: \text{"\_copula\_v\_id\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_6, \text{ARG2: } x_4), \\ h_7: \text{"exist\_q\_rel"}(\text{ARG0: } x_4, \text{RSTR: } h_8, \text{BODY: } h_9) \end{array} \right|$   
 $\left. \begin{array}{l} \{ \} \end{array} \right\rangle$

This may not actually be the most appropriate analysis for nouns which bear the =ax<sup>w</sup> enclitic. If the distribution of this enclitic is constrained to only appear on the main predicates of clauses (Zahir, 2018, 63), then it would be preferable to attach the semantics to the event-like variable contributed by the predication. Such an analysis would require the integration of the lexical rule which creates predicative nouns into an earlier stage of the morphological pipeline than I have currently. Nevertheless, because the temporal enclitic can attach to nouns (in at least some contexts), because the primary use of my grammar at the moment is parsing (as opposed to generation) and because the representation where the “now” EP attaches to the event-like variable is recoverable from the one where it attaches to the underlying noun’s individual variable (i.e. this can be unified programmatically) in principal, I allow the current representation to stand for this evaluation exercise.

### 9.1.2 Nominalizers

Lushootseed presents explicit nominalizers, treated in Beck’s materials as proclitics, which, in contrast to the subject-raising constructions discussed in Chapter 8, form gerund-like constructions. There are two types of verbal nominalizers, the so called s=nominalizer and what Beck terms the adjunctive nominalizer (ADNM=). In both cases, combinatoric results of the application of these morphemes is that the phrase refers to a reification of the underlying event — a function which is very much the converse of the predicative noun rule which builds intransitive event-like predicates from nouns. These markers posit an individual-type index over an underlying event-type variable. I have represented this along the same lines as the other contentful morphology I have described above. Like the verbal reduplication, the possessive markers and

the now enclitic, I have added a position class to the morphotactic description which contains two lexical rule types which are associated with EP contributions by the metagrammar system at customization time.

I did this, but in fact the inclusion of these nominalizers in and of itself was not helpful in increasing coverage over the development corpus. The reason for this is that in nominalized verbs, subjects are marked forms which are homophonous with the possessive affixes discussed above. Because Beck treats the nominalizers as phrase-level clitics, anything which attaches outside of nominalization must also be a clitic (under a lexical integrity hypothesis which strictly separates word-building from phrase building). The forms in my input come directly from Beck and therefore bear = symbols instead of - and in that way are not formally homophonous with the possessive affixes. Consequently, the treatment of possessive affixes described above cannot apply. Then, because such nominalized clauses have a valence requirement for at least a subject, there is always either an explicit subject marked via the periphrastic possessive or else one of these possessive clitics is present. Neither the periphrastic possessive nor the possessive clitics were implemented, and so my grammar fails to parse any sentences which include nominalized clauses.<sup>4</sup>

### 9.1.3 *Plural marker əlgʷəʔ*

Lushootseed allows the third-person subject of a clause to be marked as plural by the particle *əlgʷəʔ* (152a). The same marker can also be for objects (152b), although this usage seems uncommon to me from an informal search through the corpus. It is relatively commonly employed to mark the plurality of a third person possessor (152c).

---

<sup>4</sup>Since forking from the Grammar Matrix, libraries for possessives (Nielsen, 2018) and nominalization (Howell et al., 2018) were developed. Because these libraries were developed after my fork, I haven't yet been evaluated their utility on the Lushootseed data.

- (152) a.  $\text{?u}\check{x}^w\text{ax}^w \quad \text{al}g^w\text{a?} \quad \text{?al} \quad \text{ti?i}\check{l}$   
 $\text{?u}\check{x}^w=\text{ax}^w \quad \text{al}g^w\text{a?} \quad \text{?al} \quad \text{ti?i}\check{l}$   
 go=now PL LOC DEM

‘They go there.’

Changer (ML)

- b.  $\text{?a}\check{s}u\check{d}x^w \quad \check{c}\text{ad} \quad \text{al}g^w\text{a?}$   
 $\text{?as-}\check{s}u\check{l}\text{-dx}^w \quad \check{c}\text{ad} \quad \text{al}g^w\text{a?}$   
 STAT-SEE-DC 1SG.SBJ PL

‘I see them’

(Beck, n.d., 418)

- c.  $\text{?u?}\check{a}y\check{d}x^w \quad \check{c}\text{ad} \quad \text{ti} \quad \text{sq}^w\text{abay?}\text{s} \quad \text{al}g^w\text{a?}$   
 $\text{?u-?}\check{a}y\check{d}\text{-dx}^w \quad \check{c}\text{ad} \quad \text{ti} \quad \text{sq}^w\text{abay?}\text{-s} \quad \text{al}g^w\text{a?}$   
 PFV-find-DC 1SG.SBJ dog-3.POSS PL

‘I found their dog.’

[61]hes:1995

As with other phenomena described in this section, I targeted a subset of the possible constructions in an attempt to provide greater coverage over some grammatical low-hanging fruit. To this end I added a rudimentary treatment of the subject-marking construction where the subject is not explicit (such as (152a)). I created a new head-type, the particle (*prtcl*), and a new lexical type with no valence. Then, I defined a construction-specific rule-type in order to combine this particular lexical type ( $\text{al}g^w\text{a?}$ ) with the phrase generated by the *head-opt-subject* rule. In my grammar, the *head-subject* rule which combines a predicate with an explicit subject does not share an immediate supertype with the *head-opt-subject* rules that satisfy a subject constraint without an explicit subject. Because my goal in this rudimentary implementation was to gain some coverage without undertaking a full study of each of the combinatoric possibilities for this marker, I chose just one of the two constructions for my first pass at  $\text{al}g^w\text{a?}$ . The initial examples I looked at did not have an explicit subject, so I opted for opt and  $\text{al}g^w\text{a?}$  marking on explicit subjects is not supported by the grammar at this time. This  $\text{al}g^w\text{a?}$ -phrase is a headed-rule, so

the resulting phrase will bear the same head-features as its head-daughter, which is constrained to be generated by a *head-opt-subject* rule (153).

(153) prtcl := head.

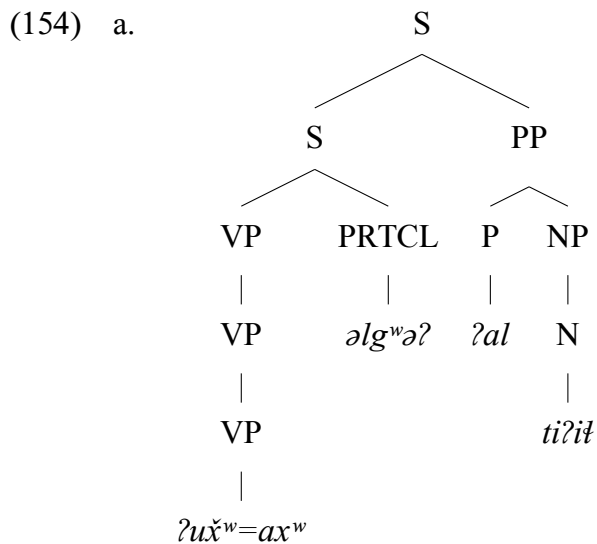
```

əlgwə?-lex := word &
 [SYNSEM.LOCAL.CAT [HEAD prtcl & [MOD < >],
 VAL [SUBJ < >,
 COMPS < >,
 SPEC < >,
 SPR < >]]].

əlgwə?-phrase := head-initial &
 [SYNSEM #synsem,
 HEAD-DTR vso-decl-head-opt-subj-phrase &
 [SYNSEM #synsem &
 [LOCAL.CONT.HOOK.XARG.PNG [PER 3rd,
 NUM pl]]],
 NON-HEAD-DTR əlgwə?-lex].

```

Adding instances of these types to the lexicon and the rules listing was enough to provide coverage for sentences such as (152a). The parse tree and semantics assigned by the grammar to this sentence is shown in (154).



- b.  $\langle h_1, e_2, \left\{ \begin{array}{l} h_3: \text{"\_?u}\check{x}^w\_v\_go\_rel"}(\text{ARG0: } e_2 \text{ ARG1: } x_4\{\text{PNG.PER } 3RD, \text{PNG.NUM } PL\}), \\ h_3: \text{"\_ax}^w\_x\_now\_rel"}(\text{ARG0: } e_2), \\ h_5: \text{"\_?al\_p\_location\_rel"}(\text{ARG0: } e_7, \text{ARG1: } e_2, \text{ARG2: } x_6), \\ h_8: \text{"\_ti?i}\check{t}\_n\_dist.dma\_rel"}(\text{ARG0: } x_6), \\ h_9: \text{"exist\_q\_rel"}(\text{ARG0: } x_6, \text{RSTR: } h_{10}, \text{BODY: } h_{11}) \\ \{\} \rangle$

#### 9.1.4 Locative PPs (*?al* phrases)

Hess (1995, 82) treats Lushootseed clausal structure as consisting of a core (the predicate and optional direct argument), an optional oblique complement, an optional temporal or locative augment and zero or more adjuncts. One relatively common type of adjunct is a phrase headed by the locative preposition *?al* (often translated as “at”). This preposition is the base for three further derived forms: *tul?al*, *dx<sup>w</sup>?al* and *li?al* often glossed as “toward”, “from” and “by way of”, respectively (155).<sup>5</sup>

- (155) a. *?ahahəx<sup>w</sup>*                      *?al ti?i}\check{t}*  
 ?<ah>a=ax<sup>w</sup>                      ?al ti?i}\check{t}  
 be.there<INTNS>=now LOC DET

‘It is there.’

Basket Ogress ML

- b. *g<sup>w</sup>əl*   *}\check{c}il*   *dx<sup>w</sup>?al*   *ti?i}\check{t} q}\check{i}lbid*  
 g<sup>w</sup>əl   }\check{c}il   dx<sup>w</sup>-?al   ti?i}\check{t} q}\check{i}lbid  
 SCONJ arrive CNTRPT-LOC DET canoe

‘Then he gets to his canoe.’

Basket Ogress (LA)

<sup>5</sup>See (Beck, n.d., 153) for further analysis and discussion of the directional prefixes.

- c. *λiʷəx<sup>w</sup> tulʔal tsiʔəʔ cədiʔ*  
*λiʷ=ax<sup>w</sup> tuʔ-ʔal tsiʔəʔ cədiʔ*  
 escape=NOW CNTRFG-LOC DET.F 3

‘He escapes from her.’

Basket Ogress ML

- d. *ʔuχ<sup>w</sup>əx<sup>w</sup> liʔʔal tiʔəʔ ləkawəkaw ʔə tiʔəʔ bək<sup>w</sup> stab*  
*ʔuχ<sup>w</sup>=ax<sup>w</sup> liʔ-ʔal tiʔəʔ lə=kaw-kaw ʔə tiʔəʔ bək<sup>w</sup> stab*  
 go=NOW PRLV-at DET PROG=DSTR-bump PR DET all what

‘She goes through there and she is bumping into everything.’

Basket Ogress (AJ)

As with my implementation of the plural marker, I sought to offer a minimally viable implementation which would aid coverage over natural text, even if it didn’t account for all possible syntactic combinations. To that end, I defined a locative preposition type which can take place in a head-modifier rule, linking the ARG1 of its predication to INDEX of the verb it attaches to. The ARG2 is the INDEX of the complement of the preposition (156).

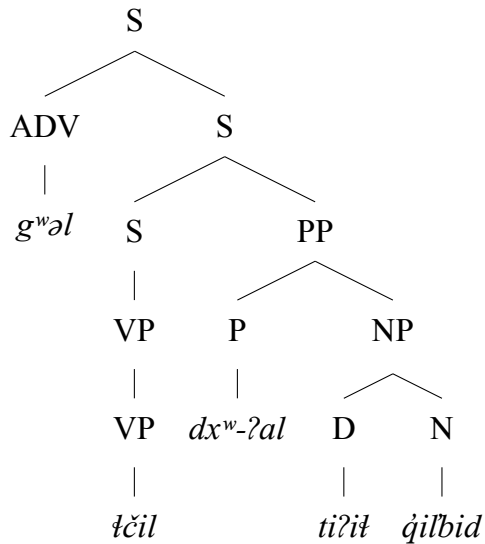
```

(156) locative-prep := basic-one-arg & norm-sem-lex-item &
 [SYNSEM.LOCAL [CAT [HEAD adp &
 [MOD < [LOCAL [CONT.HOOK.INDEX #e,
 CAT.HEAD verb]] >],
 VAL [SPR < >,
 SUBJ < >,
 COMPS < #comp >,
 SPEC < >]],
 CONT [RELS <! arg12-relation &
 [ARG0 event,
 ARG1 #e,
 ARG2 #x] !>]],
 ARG-ST < #comp &
 [LOCAL [CAT [HEAD noun,
 VAL [SPR < >,
 SPEC < >,
 SUBJ < >,
 COMPS < >]],
 CONT.HOOK.INDEX #x]] >].

```

This initial implementation only applies to verbal heads. I added an instance type to the lexicon for each of the four prepositions shown in (155) and I augmented the rule types in the word-order module to include a *head-modifier-loc-pp* rule which implements phrase-structure constraints to prevent the rule from spinning, and ensure that the preposition's valence requirements are saturated. This implementation provides the analysis shown in (157) for (155b).

(157) a.

b.  $\langle h_1, e_w,$ 

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| $h_3: \text{"_gʷəl\_sconj\_and\_rel"}(\text{ARG0: } e_5 \text{ ARG1: } h_4),$<br>$h_6: \text{"_ʈčil\_v\_arrive\_rel"}(\text{ARG0: } e_2 \text{ ARG1: } x_7),$<br>$h_8: \text{"_dxʷ-ʔal\_p\_cntrpt-location\_rel"}(\text{ARG0: } e_{10} \text{ ARG1: } e_2 \text{ ARG2: } x_9),$<br>$h_{11}: \text{"_tiʔiʈ\_q\_dist\_rel"}(\text{ARG0: } x_9 \text{ RSTR: } h_{12} \text{ BODY: } h_{13}),$<br>$h_{14}: \text{"_qilʔbid\_n\_canoe/vehicle\_rel"}(\text{ARG0: } x_9)$ |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

$\{ h_4 =_q h_6 \}$

### 9.1.5 Lexical suffixes: nouns

As mentioned in the introductory materials of Chapter 2, Lushootseed, like many other languages of the region, has a large collection of so-called lexical-suffixes which are used in creating complex words. The full syntax and semantics of Lushootseed lexical suffixes is beyond the scope of my current work, but as with other phenomena in this section, I found what I believe is a defensible way to add some coverage to the grammar without completely covering all of the possibilities. In particular for lexical suffixes, I observe that when a lexical suffix is combined with a noun, the result is still a noun and, crucially for my purposes, the valence of the base noun is unaffected. Nouns are intransitive (when they are predicative) and nouns which bear lexical suffixes may still be predicative (158a). On the other hand, the situation with verbs appears to be

more complex. Sometimes lexical suffixes fill a valence slot which would otherwise be linked to a syntactic argument, as in (158b) where the subject of the intransitive verb base *q̣<sup>w</sup>uṗ* (“be shriveled”) is provided by the lexical suffix *•ači* (“hand”). This is possible even with valence increasing affixes which attach outside the lexical suffix, as shown in (158c). In some other cases, the lexical suffix seems to provide what should be a syntactic argument which would require a valence change to the base, but there’s no marking at all to indicate this (158d). In still other cases, the lexical suffix seems to be able to affect the base-verb’s meaning without affecting the valence (158e).

- (158) a. *q<sup>w</sup>i?ad k<sup>w</sup>ədi? di?ucid*  
*q<sup>w</sup>i?ad k<sup>w</sup>ədi? di?•ucid*  
 call.out DET other.side•mouth  
 ‘The one across the river calls out.’ Coyote and his daughter (ML)
- b. *puṭax<sup>w</sup> ?asq̣<sup>w</sup>uṗq̣<sup>w</sup>uṗači?*  
*put=ax<sup>w</sup> ?as-q̣<sup>w</sup>uṗ-q̣<sup>w</sup>uṗ•ači?*  
 really=now STAT-DSTR-shriveled•hand  
 ‘(His) hands are all shriveled up.’ Black Bear and Fish Hawk (Edward Sam)
- c. *yəxi tubəhudabactəbax<sup>w</sup>*  
*yəxi tu=bə=hud•abac-t-b=ax<sup>w</sup>*  
 because PST=add=burn•body-ICS-PASS=now  
 ‘because her body had been burned.’ Basket Ogress (ML)
- d. *ṣ<sup>w</sup>uṭ ?uq̣cač* *ti?iṭ sčətx<sup>w</sup>əd*  
*ṣ<sup>w</sup>uṭ ?u-q̣c•ač* *ti?iṭ sčətx<sup>w</sup>əd*  
 only PFV-strike.back.of.head•head DET bear  
 ‘Black Bear strikes the back of his head.’ Black Bear and Fish Hawk (ES)

- e. *huy* *ʔahəx<sup>w</sup>* *ʔəspədič*  
*huy* *ʔa=ax<sup>w</sup>* *ʔas-pəd•ič*  
 sconj be.there=now stat-buried•covering

‘Then she is buried right there.’

Basket Ogress (ML)

Given that the noun attachments leave the morphosyntactic properties unchanged, I chose to list the nouns which bear lexical suffixes directly in the lexicon with the lexical suffix already incorporated. In this way, I gain some lexical coverage over these forms using the noun-classes I’ve already defined. For verbs, due to the morphotactic variation I observed in a few readily accessible example, I chose to leave them unanalyzed.

## 9.2 Results

In this section I describe the results of the exercise. Details about the grammar’s performance on each of the four stories are provided in separate subsections below. In the first subsection I describe the data preparation I did on the raw texts in order to prepare them for use in the exercise.

### 9.2.1 Data Preparation

While the data for this evaluation were first published in Hess (1995) as Lushootseed-only texts, the actual input for this evaluation is based on the morpheme-segmented line from these stories as published in Beck and Hess (2014). There are a few notable differences, however. All of these stories contain songs. Some of the songs have linguistic lyrics which the grammar may be able to parse, such as the principle refrain of *ʔaʔačapəd*’s (“Ant”) song. Beck and Hess provide full IGT for the songs just like the spoken text, and indicate that they are sung by surrounding the free-translation line with eighth-note symbols (159).

(159) *łaxil g<sup>w</sup>əl bəlaxil*  
 łax-il g<sup>w</sup>əl bə=ləx-il  
 dark-inch sconj add=light-inch

♪ “Night and again day,” ♪

Bear and Ant (ES)

However, there are other parts of the songs which are melody-only, or, at the very least, the grammar wouldn’t be the right tool for analyzing these lines. These lines are included in the original published in LR1, and they also appear in the Beck and Hess corpus as full IGT (160). I prepared the data for the evaluation by removing melody-lines such as (160). I kept, however, the lines from songs with linguistic lyrics such as (159).

(160) *we<sup>r</sup> x<sup>w</sup>e?e<sup>r</sup>? e<sup>r</sup> we<sup>r</sup> x<sup>w</sup>e?e<sup>r</sup>?*  
*we<sup>r</sup> x<sup>w</sup>e?e<sup>r</sup>? e<sup>r</sup> we<sup>r</sup> x<sup>w</sup>e?e<sup>r</sup>?*  
*we<sup>r</sup> x<sup>w</sup>e?e<sup>r</sup>? e<sup>r</sup> we<sup>r</sup> x<sup>w</sup>e?e<sup>r</sup>?*

“weee x<sup>w</sup>e?eee? eee weee x<sup>w</sup>e?eee?”

Bear and Ant (ES)

Another aspect of these texts which is all the more salient because of the prominence of songs, is that they contain duplicated lines. Duplicate lines are interesting in an evaluation like this because, on the one hand, these are natural texts so if a sentence-type is repeated, it would be all-the-more useful for a Lushootseed processing system to handle it. Among genre of natural language, however, songs and poems arguably contain even more repetition than most. Because my grammar encodes a strict notion of grammaticality and songs and poems form a domain where the syntactic rules which apply to speech are often stretched (including with respect to repetition), there’s an argument to be made that evaluating a system built to handle standard speech syntax against song-lyrics is unfair. In this regard, I chose to present the results with duplicates removed (“deduped”) alongside the results with them left in. Now I can turn to the results tables themselves.

### 9.2.2 *Bear and Ant (Edward Sam)*

Of the four stories used in this evaluation, this one has the most repeated lines and therefore the greatest difference between the original and deduped profiles. Because the grammar does not parse the main refrain of Ant’s song (159), the results on the deduped profile show a correspondingly marked improvement in coverage. The results for the quadrant (System I and II  $\times$  original and deduped) are presented in Table 9.1. System I (“everything is a predicate”) and System II (“restricted predication contexts”) provided identical coverage over the text of this story (33.3, 46.7% (original, deduped)). However, the metric of note in Table 9.1 is the column labelled “distinct readings”. System I was notably more ambiguous than System II in getting to this coverage, providing averages of 1.59 and 1.61 readings per sentence parsed, while System II maintains a 1:1 ratio between sentences parsed and proffered readings for those sentences. Error analysis reveals that the additional readings from System I are due to the underconstrained SPEC value on the System I version of the predicative noun lexical rule. This leads to incorrect parses of certain determiner-noun sequences.

Specifically, Lushootseed has determiners which correspond to homophonous demonstrative pronouns. From the perspective of lexical parsing then, the sequence *determiner-noun* is homophonous with a corresponding sequence *demonstrative-noun*. Correct grammars would rule out the *demonstrative-noun* sequence at the syntactic level, because if the first element is a demonstrative pronoun, the bare noun is left out of the spanning parse. In the problematic parses coming from System I, the predicative noun rule fires, extracting the subject and leaving the specifier value unconstrained. Then, in certain cases, the grammar finds an incorrect parse where a leftover noun combines with a VP or even an S constituent in a head-specifier phrase. Such a parse, with the S providing a specifier for a noun, is incorrect. While further constraints on System I could presumably correct this error while maintaining the desirable features of System I,<sup>6</sup> the purpose of this evaluation is to take stock of things as they stand. So I have chosen to

---

<sup>6</sup>The trivial solution, to constrain the specifier value on the output of the predicative noun lexical rule, in naive form, defeats some of the purposes of System I: constraining it to the empty list would prevent predicative nouns from appearing inside determiner phrases (a feature of System II), and constraining it to pass up the noun’s

| <b>'Bear and Ant: original : System I' Coverage Profile</b> |                    |                    |                      |                          |                      |                         |
|-------------------------------------------------------------|--------------------|--------------------|----------------------|--------------------------|----------------------|-------------------------|
| <b>Aggregate</b>                                            | <b>total items</b> | <b>word string</b> | <b>lexical items</b> | <b>distinct analyses</b> | <b>total results</b> | <b>overall coverage</b> |
|                                                             | #                  | $\phi$             | $\phi$               | $\phi$                   | #                    | %                       |
| $10 \leq i\text{-length} < 15$                              | 3                  | 10.33              | 14.00                | 0.00                     | 0                    | 0.0                     |
| $5 \leq i\text{-length} < 10$                               | 19                 | 6.05               | 7.50                 | 3.25                     | 4                    | 21.1                    |
| $0 \leq i\text{-length} < 5$                                | 65                 | 2.80               | 3.19                 | 1.32                     | 25                   | 38.5                    |
| <b>Total</b>                                                | <b>87</b>          | <b>3.77</b>        | <b>4.14</b>          | <b>1.59</b>              | <b>29</b>            | <b>33.3</b>             |
| <b>'Bear and Ant: deduped: System I' Coverage Profile</b>   |                    |                    |                      |                          |                      |                         |
| $10 \leq i\text{-length} < 15$                              | 3                  | 10.33              | 14.00                | 0.00                     | 0                    | 0.0                     |
| $5 \leq i\text{-length} < 10$                               | 18                 | 6.06               | 7.55                 | 3.25                     | 4                    | 22.2                    |
| $0 \leq i\text{-length} < 5$                                | 39                 | 2.69               | 3.37                 | 1.33                     | 24                   | 61.5                    |
| <b>Total</b>                                                | <b>60</b>          | <b>4.08</b>        | <b>4.82</b>          | <b>1.61</b>              | <b>28</b>            | <b>46.7</b>             |
| <b>'Bear and Ant: original: System II' Coverage Profile</b> |                    |                    |                      |                          |                      |                         |
| $10 \leq i\text{-length} < 15$                              | 3                  | 10.33              | 14.00                | 0.00                     | 0                    | 0.0                     |
| $5 \leq i\text{-length} < 10$                               | 19                 | 6.05               | 7.60                 | 1.00                     | 4                    | 21.1                    |
| $0 \leq i\text{-length} < 5$                                | 65                 | 2.80               | 3.19                 | 1.00                     | 25                   | 38.5                    |
| <b>Total</b>                                                | <b>87</b>          | <b>3.77</b>        | <b>4.05</b>          | <b>1.00</b>              | <b>29</b>            | <b>33.3</b>             |
| <b>'Bear and Ant: deduped : System II' Coverage Profile</b> |                    |                    |                      |                          |                      |                         |
| $10 \leq i\text{-length} < 15$                              | 3                  | 10.33              | 14.00                | 0.00                     | 0                    | 0.0                     |
| $5 \leq i\text{-length} < 10$                               | 18                 | 6.06               | 7.60                 | 1.00                     | 4                    | 22.2                    |
| $0 \leq i\text{-length} < 5$                                | 39                 | 2.69               | 3.37                 | 1.00                     | 24                   | 61.5                    |
| <b>Total</b>                                                | <b>60</b>          | <b>4.08</b>        | <b>4.76</b>          | <b>1.00</b>              | <b>28</b>            | <b>46.7</b>             |

Table 9.1: The evaluation results for Bear and Ant

continue the evaluation without any reworking of System I. In each of the cases where System I was providing more than one parse for a particular item which System II analyzes with a single reading, a sensible reading was also found. So the extra ambiguity of System I can be seen as a kind of overgeneration.

---

specifier requirement would prevent it from analyzing most of the test suite of Chapter 8, where the matrix-clause predicative nouns do not bear determiners.

### 9.2.3 *Bear and Fishhawk (Edward Sam)*

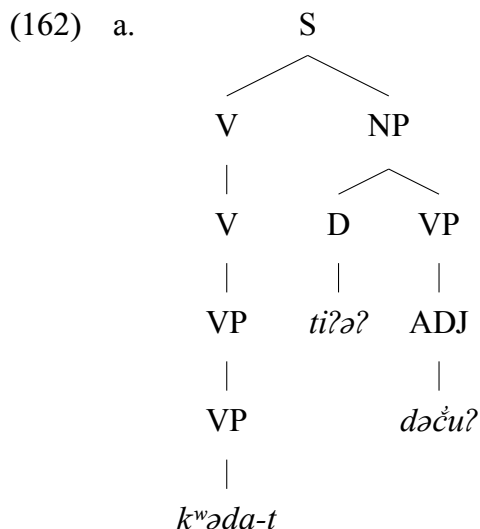
The results for Bear and Fishhawk are summarized in Table 9.2. This time, System I provides just a bit more coverage than System II, with five more sentences parsed. As was the case with the Bear and Ant story, System I is more ambiguous, for the same reason discussed above (the underconstrained SPEC values).

There are six items which are analyzed by System I, but not by System II. Of these, in three cases, all of the readings are incorrect, the error being ultimately related to the underconstrained SPEC value. However, in a third case, System I is able to provide a parse with a plausibly correct reading for the sentence in (161). The parse tree and the semantics produced are given in (162).

(161) *kʷədad tiʔəʔ dəčʰuʔ*  
*kʷəda-t tiʔəʔ dəčʰuʔ*  
 taken-ICS DET one

‘He takes one.’

Bear and Fishhawk (ES)



- b.  $\langle h_1, e_2 \{SF \textit{ PROP-OR-QUES}, E.ASPECT \textit{ IMPERFECTIVE}\},$   
 $\left. \begin{array}{l} h_3: \textit{ "_k^w\text{ə}da\_v\_take/grab\_rel"}(ARG0: e_2, ARG1: x_4), \\ h_3: \textit{ "_t\_x\_ics\_rel"}(ARG0: e_2, ARG1: x_5), \\ h_6: \textit{ "_ti?ə?_q\_prox\_rel"}(ARG0: x_4, RSTR: h_7, BODY: h_8), \\ h_9: \textit{ "_d\text{ə}č\text{u}?\_j\_one\_rel"}(ARG0: e_{10}, ARG1: x_4) \\ \{ h_7 =_q h_9 \} \end{array} \right\rangle$

In the other two cases, System I provides coverage where System II fails for one of the same reasons discussed in Chapter 8, System I is able to cover nouns with general prefixes, even in embedded contexts. In (163a), the additive general prefix appears on the noun *sčx<sup>w</sup>əs* (“fat/grease”) indicting the repeating of the episode where Fishhawk magically produces oil by rubbing his hands together. System I is able to parse this because the predicative noun rule is able to fire in embedded contexts (in fact, under System I it must fire in order for the noun to combine with a determiner), leaving an event-like index variable that is compatible with the additive semantics (which specifies the property [MOOD *add*] on its index). Similarly, in (163b), the subjunctive marker *g<sup>w</sup>ə=* appears on an embedded nominal.<sup>7</sup>

- (163) a. *g<sup>w</sup>əl huy ?u?abyidəx<sup>w</sup> ti?ił sčətx<sup>w</sup>əd* [*?ə ti?ił bəsčx<sup>w</sup>əs*]  
*g<sup>w</sup>əl huy ?u-?ab-yi-t=ax<sup>w</sup> ti?ił sčətx<sup>w</sup>əd ?ə ti?ił bə=sčx<sup>w</sup>əs*  
 SCONJ SCONJ PFV-extend-DAT-ICS=NOW DET bear PR DET ADD=fat

‘And next he gives Black Bear more fat,’

Bear and Fishhawk (ES)

<sup>7</sup>In fact, subjunctive marking is grammatically required on the complement of the negative verb *x<sup>w</sup>i?*. A more complete approach to negative constructions involving *x<sup>w</sup>i?* would most-likely have *x<sup>w</sup>i?* constrain it’s complement’s index’s MOOD value to subjunctive. An interesting ramification of this is that the determiner phrase would need to expose the event-like variable instead of or alongside the instance variable associated with the underlying noun. To me, the matter of how to correctly specify Lushootseed variable types in syntactic indexes is one of the most important topics of further study which is highlighted by the case study of Chapter 8.

- b.  $x^w i ? \partial x^w$  [ $k^w i$ ]  $g^w \partial s \check{x}^w \partial s$   
 $x^w i ? = a x^w$   $k^w i$   $g^w \partial = s \check{x}^w \partial s$   
 NEG=NOW DET SBJV=fat

‘There is no fat.’

Bear and Fishhawk (ES)

System II, however, also provided coverage for one item which System I did not. This three-word item (164) actually has some tricky syntax. Intuitively, a correct parse of this item would have an optional subject for the negative verb  $x^w i ? \partial x^w$  meaning “there is no X” where X is represented by the optional subject. Then, that entire phrase would be the complement of  $? \partial s a y d x^w \partial \check{x}^w$  (“knows”), a transitive verb with a subject which is the knower and the object which is the thing known. This phrase is not the grammatical object of the matrix verb  $b \partial \check{s} u d x^w$  (“see again”), but is instead a modifier clause (at least according to the translation). Since my grammar has no way to analyze a clause as a modifier, System II did not come up with a plausible analysis, but one where the phrase which might be glossed “knows there are none left” is the thing which was seen again.

- (164)  $b \partial \check{s} u d x^w$   $? \partial s a y d x^w \partial \check{x}^w$   $x^w i ? \partial x^w$   
 $b \partial = \check{s} u \check{t} - d x^w$   $? a s - h a y - d x^w = a x^w$   $x^w i ? = a x^w$   
 ADD=SEE-DC STAT-KNOWN-DC=NOW NEG=NOW

‘He looks again thinking there are none left.’

Bear and Fishhawk (ES)

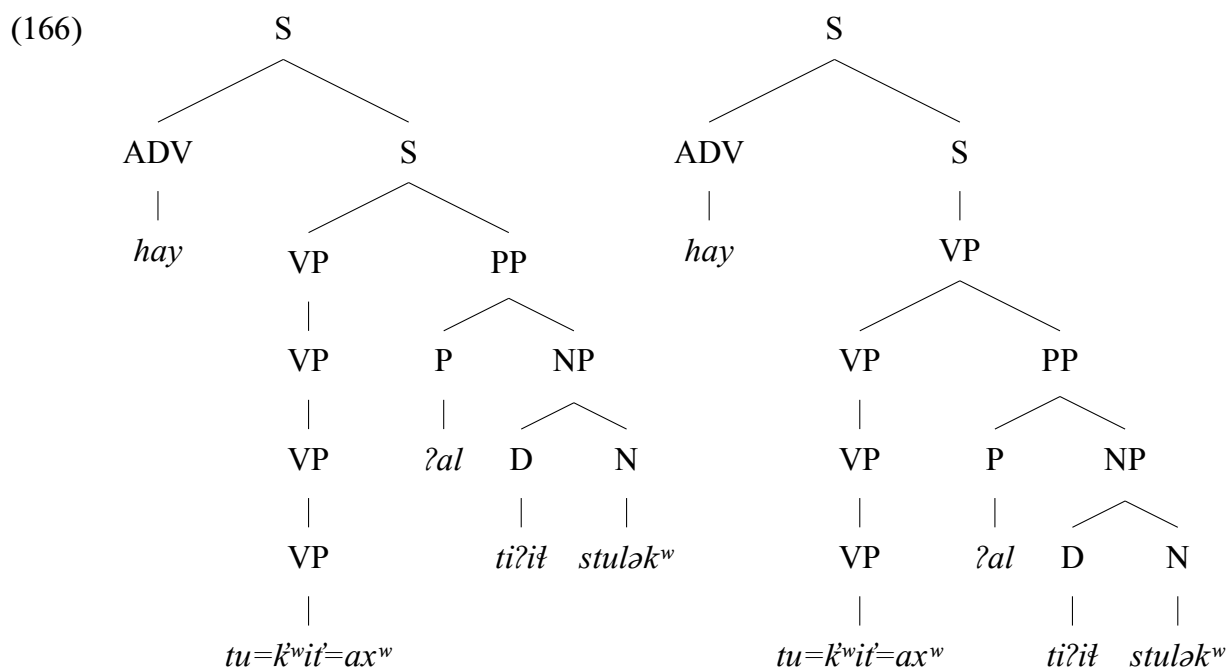
Also of note in comparison to the previously discussed results: while System II still produces less ambiguity than System I (and it’s still the case that most of the ambiguity produced by System I represents an error), on this text System II also provided some ambiguity. There were three sentences in the story to which System II assigns more than one parse. The first of these is a textbook example of spurious syntactic ambiguity: two parse trees which yield the same semantic representation due to the application of the same syntactic rules in more than one order. The sentence in (165) employs a locative modifier phrase, so my implementation of

this construction (discussed above in §9.1.4) fires. However, nothing prevents the head-modifier rule from applying before or after the head-complement rule (166), so the System II grammar finds two parses with equivalent semantics.<sup>8</sup>

- (165) *hay tuk<sup>w</sup>it<sup>ʔ</sup>ax<sup>w</sup>                    ʔal tiʔit<sup>ʔ</sup> stulək<sup>w</sup>*  
*hay tu=k<sup>w</sup>it<sup>ʔ</sup>=ax<sup>w</sup>                    ʔal tiʔit<sup>ʔ</sup> stulək<sup>w</sup>*  
 SCONJ PST=shoreward=now LOC DET river

‘So (Fish Hawk) went down to the river.’

Bear and Fishhawk (ES)



The second ambiguous sentence from System II is also erroneous. In this case, the grammar finds three readings for (167). In fact, neither of the two systems is expected to parse the appositive construction found in this example. The erroneous parses are due to an error in the implementation of the allative applicative (ALTV) lexical rule. This rule correctly links the subject’s semantics into the predication provided by the verbal base, but it fails to constrain the

<sup>8</sup>The System I grammar finds this attachment ambiguity as well, in addition to other, erroneous parses due to the same issues with System I discussed above.

syntactic features of the subject sufficiently. Because of that, all three parses have analyzed the form *ćixćix* (“fish hawk”) as if it were the applicative argument: “his companion arrived at fish hawk”. The three analyses correspond to three treatments of the noun: the bare noun, the predicative noun, and the sentential noun (predicative noun followed by the head-optional subject rule).

- (167) *g<sup>w</sup>əl* *čisəb*                    *ʔə* *tiʔiʔ* *ʔayəds*                    *ćixćix*  
*g<sup>w</sup>əl* *čil-s-b*                    *ʔə* *tiʔiʔ* *ʔayəd-s*                    *ćixćix*  
 SCONJ arrive-ALTV-PASS PR DIST companion-3.POSS fish.hawk

‘And he is arrived at by his friend Fish Hawk.’

Bear and Fishhawk (ES)

The six parses the System II grammar finds for the third item correspond to both of the just-described error conditions being present in the same sentence in addition to the incorrect lexical analysis of a *determiner-noun* sequence as *demonstrative-noun* (168). As in (167), the underconstraint of the syntactic category of a morphologically added valence item allows a noun to fill a syntactic valence slot which it really shouldn’t. As in (165), both the high and low application of the locative modifier phrase create pairs of syntactic trees with equivalent readings. In this case, however, unlike (167), the relevant NP construction does not contain an appositive (a phenomenon which is unaddressed in my grammar) but instead just a determiner-noun, so among the six parses are two with the right semantics.

- (168) *g<sup>w</sup>əl* *ʔuχ<sup>w</sup>tubəx<sup>w</sup>*                    *dx<sup>w</sup>ʔal*                    *tiʔiʔ* *ʔalʔals*  
*g<sup>w</sup>əl* *ʔuχ<sup>w</sup>-tx<sup>w</sup>-b=ax<sup>w</sup>*                    *dx<sup>w</sup>-ʔal*                    *tiʔiʔ* *ʔalʔal-s*  
 SCONJ go-ECS-PASS=NOW CNTRPT-LOC DET house-3.POSS

‘And he is taken to his house.’

Bear and Fishhawk (ES)

In all, the added ambiguity represents errors in System II. Despite this, on the whole, System II outperformed System I on this test because apart from the two sentences presented above, the added coverage of System I was all error.

| <b>'Bear and Fishhawk: original: System I' Coverage Profile</b>  |                    |                    |                      |                          |                      |                         |
|------------------------------------------------------------------|--------------------|--------------------|----------------------|--------------------------|----------------------|-------------------------|
| <b>Aggregate</b>                                                 | <b>total items</b> | <b>word string</b> | <b>lexical items</b> | <b>distinct analyses</b> | <b>total results</b> | <b>overall coverage</b> |
|                                                                  | #                  | $\phi$             | $\phi$               | $\phi$                   | #                    | %                       |
| $15 \leq i\text{-length} < 20$                                   | 1                  | 18.00              | 0.00                 | 0.00                     | 0                    | 0.0                     |
| $10 \leq i\text{-length} < 15$                                   | 1                  | 10.00              | 0.00                 | 0.00                     | 0                    | 0.0                     |
| $5 \leq i\text{-length} < 10$                                    | 38                 | 6.13               | 8.95                 | 3.86                     | 14                   | 36.8                    |
| $0 \leq i\text{-length} < 5$                                     | 57                 | 2.65               | 3.49                 | 1.66                     | 32                   | 56.1                    |
| <b>Total</b>                                                     | <b>97</b>          | <b>4.25</b>        | <b>5.34</b>          | <b>2.33</b>              | <b>46</b>            | <b>47.4</b>             |
| <b>'Bear and Fishhawk: deduped: System I' Coverage Profile</b>   |                    |                    |                      |                          |                      |                         |
| $15 \leq i\text{-length} < 20$                                   | 1                  | 18.00              | 0.00                 | 0.00                     | 0                    | 0.0                     |
| $10 \leq i\text{-length} < 15$                                   | 1                  | 11.00              | 0.00                 | 0.00                     | 0                    | 0.0                     |
| $5 \leq i\text{-length} < 10$                                    | 37                 | 6.16               | 8.95                 | 3.86                     | 14                   | 37.8                    |
| $0 \leq i\text{-length} < 5$                                     | 56                 | 2.62               | 3.49                 | 1.66                     | 32                   | 57.1                    |
| <b>Total</b>                                                     | <b>95</b>          | <b>4.24</b>        | <b>5.34</b>          | <b>2.33</b>              | <b>46</b>            | <b>48.4</b>             |
| <b>'Bear and Fishhawk: original: System II' Coverage Profile</b> |                    |                    |                      |                          |                      |                         |
| $15 \leq i\text{-length} < 20$                                   | 1                  | 18.00              | 0.00                 | 0.00                     | 0                    | 0.0                     |
| $10 \leq i\text{-length} < 15$                                   | 1                  | 10.00              | 0.00                 | 0.00                     | 0                    | 0.0                     |
| $5 \leq i\text{-length} < 10$                                    | 38                 | 6.13               | 9.00                 | 1.80                     | 10                   | 26.3                    |
| $0 \leq i\text{-length} < 5$                                     | 57                 | 2.65               | 3.49                 | 1.10                     | 31                   | 54.4                    |
| <b>Total</b>                                                     | <b>97</b>          | <b>4.25</b>        | <b>5.29</b>          | <b>1.27</b>              | <b>41</b>            | <b>42.3</b>             |
| <b>'Bear and Fishhawk: deduped: System II' Coverage Profile</b>  |                    |                    |                      |                          |                      |                         |
| $15 \leq i\text{-length} < 20$                                   | 1                  | 18.00              | 0.00                 | 0.00                     | 0                    | 0.0                     |
| $10 \leq i\text{-length} < 15$                                   | 1                  | 10.00              | 0.00                 | 0.00                     | 0                    | 0.0                     |
| $5 \leq i\text{-length} < 10$                                    | 37                 | 6.16               | 9.00                 | 1.80                     | 10                   | 27.0                    |
| $0 \leq i\text{-length} < 5$                                     | 56                 | 2.62               | 3.49                 | 1.10                     | 31                   | 55.4                    |
| <b>Total</b>                                                     | <b>95</b>          | <b>4.24</b>        | <b>5.29</b>          | <b>1.27</b>              | <b>41</b>            | <b>43.2</b>             |

Table 9.2: The evaluation results for Bear and Fishhawk

#### 9.2.4 Coyote and Big Rock (Edward Sam)

For the third story in this evaluation, System I provides parses for three more items than System II and again results in a higher rate of ambiguity. The results of running the evaluation on the profiles for *Bear and Fishhawk* are summarized in Table 9.3. The first of the three additionally analyzed items is shown in (169). System II cannot accommodate this item because *ǰʷul* (“only”) is listed as a scopal adverb in my grammar, looking for a verbal head, and *ǰaʎaʎ* (“refuse/worthless stuff”) is listed as a noun.<sup>9</sup> System I, on the other hand, because of the head-type change of the predicative noun rule (recall that under System II, predicative nouns retain their head value but under System I, they become verbal), is able to analyze the noun phrase *tiʔit ǰʷul ǰaʎaʎ* due to this conversion. Although the System I grammar is able to get the syntax plausibly correct, the semantics is ill-formed (170). The problem is that the scopal-adjective, which wants to take its modificand’s scopal handle as its ARG1, is grabbing an underconstrained handle after the application of the predicative noun rule. That this represents an oversight in the implementation of the predicative noun rule is shown by the fact that the other scopal adverb in the sentence, the sentential conjunction *huy*, properly accesses its argument’s local top handle, using the same scopal modifier rule to combine.

---

<sup>9</sup>Beck (p.c.) points out that a better analysis for *ǰaʎaʎ* would be to treat it as a verb, not a noun. In the materials I was working with, the form never appears with a stative marker (ensuring that it must be a verb) nor with a possessive (ensuring that it must be a noun). However, the example in (1), where *ǰaʎaʎ* appears as an attributive adjective, supports Beck’s claim.

- (i) *tuǰʷ čəʔ ǰaʎaʎ tukʷədalikʷ tuʎʔa*  
*tuǰʷ čəʔ ǰaʎaʎ tu=kʷəda-alikʷ tuʎʔa*  
 just 1PL.SBJ worthless PST=taken-ACT CNTRFG-be.there

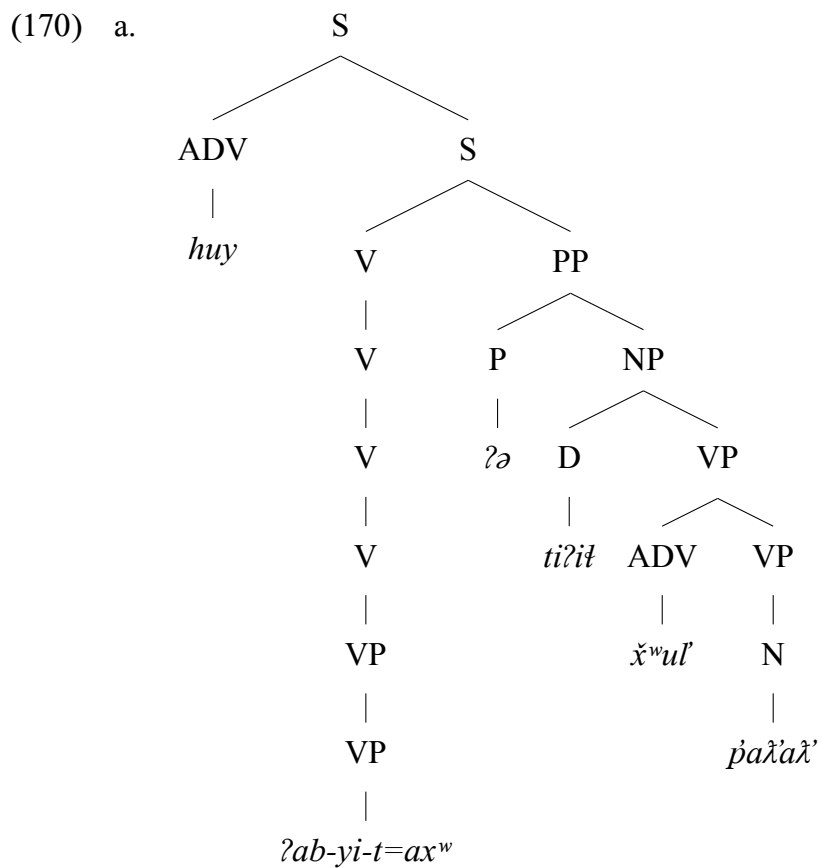
‘We just took worthless (bits) from there.’

Changer (ML)

- (169) *huy* *ʔabyidəx<sup>w</sup>*                    *ʔə* *tiʔiʔ* *ʃ<sup>w</sup>ul'* *paʔaʔ*  
*huy* *ʔab-yi-t=ax<sup>w</sup>*                    *ʔə* *tiʔiʔ* *ʃ<sup>w</sup>ul'* *paʔaʔ*  
 SCONJ extend-DAT-ICS=NOW PR DIST only worthless

‘Then, he gives him just what is worthless.’

Coyote and Big Rock (ES)



- b.  $\langle h_1, e_2 \{SF \text{ PROP-OR-QUES, E.ASPECT IMPERFECTIVE}\},$
- |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| $h_3: \text{"\_huy\_sconj\_then\_rel"}(\text{ARG0: } e_5, \text{ARG1: } h_4),$<br>$h_6: \text{"\_?ab\_extend\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_7),$<br>$h_6: \text{"\_yid\_x\_dat\_rel"}(\text{ARG0: } e_2, \text{ARG1: } i_8),$<br>$h_6: \text{"\_ics\_x\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_9),$<br>$h_{10}: \text{"\_ax}^w \text{\_x\_now\_rel"}(\text{ARG0: } e_2),$<br>$h_{11}: \text{"\_ti?i\text{†}\_q\_dist\_rel"}(\text{ARG0: } x_7, \text{RSTR: } h_{12}, \text{BODY: } h_{13}),$<br>$h_{14}: \text{"\_x}^w \text{\_u\text{†}\_r\_only\_rel"}(\text{ARG0: } e_{16}, \text{ARG1: } h_{15}),$<br>$h_{17}: \text{"\_p\text{†}\_a\text{†}\_a\text{†}\_n\_trash/worthless.stuff\_rel"}(\text{ARG0: } x_{18} \{PNG.PER 3RD\}),$<br>$h_{19}: \text{"\_copula\_v\_id\_rel"}(\text{ARG0: } e_{20}, \text{ARG1: } x_7, \text{ARG2: } x_{18}),$<br>$h_{21}: \text{"\_exist\_q\_rel"}(\text{ARG0: } x_{18}, \text{RSTR: } h_{22}, \text{BODY: } h_{23})$<br>$\{ h_4 =_q h_6 \}$ |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

The second item parsed by System I but not System II is, from the perspective of error analysis, isomorphic to (169). In (171), *hik<sup>w</sup>* (“big”) is listed as an adverb, and therefore can’t combine directly with *čłá?* (“stone”) under System II.<sup>10</sup> Then, as before, a reasonable syntactic analysis is found among the results of System I, but the failings of the predicative noun rule just mentioned prevent the adverb from linking into the representation correctly.

<sup>10</sup>Beck (p.c.) suggests that *hik<sup>w</sup>* might be better treated as a verb or adjective. Indeed, such a treatment works well for the example under discussion. My inclusion of it as an adverb is based on sentences such as (1), where, I think, an adverbial treatment is reasonable.

- (i) *hi<sup>k</sup>w čł*      *?uhəli?dub*  
*hik<sup>w</sup> čł*      *?u-həli?-dx<sup>w</sup>-b*  
 big    1PL.SBJ PFV-alive-DC-PASS

‘Greatly have we been aided.’

Pheasant and Raven (ML)

- (171) *g<sup>w</sup>əl šudx<sup>w</sup>əx<sup>w</sup> tiʔiʔ hik<sup>w</sup> čłaʔ*  
*g<sup>w</sup>əl šuʔ-dx<sup>w</sup>=ax<sup>w</sup> tiʔiʔ hik<sup>w</sup> čłaʔ*  
 SCONJ see-DC=NOW DET big stone

‘And he sees Big Rock.’

Coyote and Big Rock (ES)

The final additional parse from System I is erroneous. In this case, an intransitive verb takes an oblique adjunct, a construction as of yet unimplemented in my grammar (172). The errant parse from System I is due to one of the error types discussed above: the underconstrained SPEC value of the predicative noun rule allows an erroneous analysis whereby the sequence determiner-noun is treated as demonstrative-noun and the stranded noun is reattached via a head-specifier rule.

- (172) *dəg<sup>w</sup>ag<sup>w</sup>iləx<sup>w</sup> ʔə tiʔiʔ [s]əsliʔluʔ*  
*dək<sup>w</sup>a-ag<sup>w</sup>il=ax<sup>w</sup> ʔə tiʔiʔ səsliʔluʔ*  
 inside-AUTO=NOW PR DET little.cave

‘He gets inside a small cave.’

Coyote and Big Rock (ES)

I think the takeaway from these three parses is that if I had correctly constrained the local top handle of the noun in the predicative noun rule, System I (by virtue of its casting noun into verb in its version of the rule) would have provided correct additional coverage for two sentences which are not analyzable under System II.

The other bit to look at here is the ambiguous sentences under System II. There were three items which provided more than one parse. Two of these were immediately associable with one of the error classes above for reasons discussed above, where the underconstrained head-type of the morphologically added valence item leads to a cascade of incorrect rule combinations. The third item, however (173) presented a surprisingly shaped additional tree whereby the initial interjection, the sentential adverb, and the subject marker were combined into a constituent (174).



| <b>'Coyote and Big Rock: original: System I' Coverage Profile</b>  |                              |                                              |                                                |                                                    |                                |                                   |
|--------------------------------------------------------------------|------------------------------|----------------------------------------------|------------------------------------------------|----------------------------------------------------|--------------------------------|-----------------------------------|
| <b>Aggregate</b>                                                   | <b>total<br/>items<br/>#</b> | <b>word<br/>string<br/><math>\phi</math></b> | <b>lexical<br/>items<br/><math>\phi</math></b> | <b>distinct<br/>analyses<br/><math>\phi</math></b> | <b>total<br/>results<br/>#</b> | <b>overall<br/>coverage<br/>%</b> |
| $10 \leq i\text{-length} < 15$                                     | 3                            | 10.67                                        | 14.00                                          | 0.00                                               | 0                              | 0.0                               |
| $5 \leq i\text{-length} < 10$                                      | 32                           | 6.03                                         | 8.12                                           | 3.64                                               | 11                             | 34.4                              |
| $0 \leq i\text{-length} < 5$                                       | 45                           | 3.04                                         | 3.75                                           | 1.75                                               | 24                             | 53.3                              |
| <b>Total</b>                                                       | <b>80</b>                    | <b>4.53</b>                                  | <b>5.53</b>                                    | <b>2.34</b>                                        | <b>35</b>                      | <b>43.7</b>                       |
| <b>'Coyote and Big Rock: deduped: System I' Coverage Profile</b>   |                              |                                              |                                                |                                                    |                                |                                   |
| $10 \leq i\text{-length} < 15$                                     | 3                            | 10.67                                        | 14.00                                          | 0.00                                               | 0                              | 0.0                               |
| $5 \leq i\text{-length} < 10$                                      | 31                           | 6.00                                         | 8.12                                           | 3.64                                               | 11                             | 35.5                              |
| $0 \leq i\text{-length} < 5$                                       | 42                           | 3.07                                         | 3.81                                           | 1.78                                               | 22                             | 54.8                              |
| <b>Total</b>                                                       | <b>76</b>                    | <b>4.57</b>                                  | <b>5.61</b>                                    | <b>2.38</b>                                        | <b>34</b>                      | <b>44.7</b>                       |
| <b>'Coyote and Big Rock: original: System II' Coverage Profile</b> |                              |                                              |                                                |                                                    |                                |                                   |
| $10 \leq i\text{-length} < 15$                                     | 3                            | 10.67                                        | 14.00                                          | 0.00                                               | 0                              | 0.0                               |
| $5 \leq i\text{-length} < 10$                                      | 32                           | 6.03                                         | 8.20                                           | 1.78                                               | 9                              | 28.1                              |
| $0 \leq i\text{-length} < 5$                                       | 45                           | 3.04                                         | 3.70                                           | 1.04                                               | 23                             | 51.1                              |
| <b>Total</b>                                                       | <b>80</b>                    | <b>4.53</b>                                  | <b>5.51</b>                                    | <b>1.25</b>                                        | <b>32</b>                      | <b>40.0</b>                       |
| <b>'Coyote and Big Rock: deduped: System II' Coverage Profile</b>  |                              |                                              |                                                |                                                    |                                |                                   |
| $10 \leq i\text{-length} < 15$                                     | 3                            | 10.67                                        | 14.00                                          | 0.00                                               | 0                              | 0.0                               |
| $5 \leq i\text{-length} < 10$                                      | 31                           | 6.00                                         | 8.20                                           | 1.78                                               | 9                              | 29.0                              |
| $0 \leq i\text{-length} < 5$                                       | 42                           | 3.07                                         | 3.77                                           | 1.05                                               | 22                             | 52.4                              |
| <b>Total</b>                                                       | <b>76</b>                    | <b>4.57</b>                                  | <b>5.60</b>                                    | <b>1.26</b>                                        | <b>31</b>                      | <b>40.8</b>                       |

Table 9.3: The evaluation results for Coyote and Big Rock

The numbers for the evaluation on “Mink” are summarized in Table 9.4. Although the two systems provided equivalent coverage, error analysis of the results highlighted a new error type which was available in both System I and System II (175).

- (175) *huy* *fičitəbəx<sup>w</sup>* *ti?ił sčali? ?ə ti?ił čx<sup>w</sup>əlu?*  
*huy* *fiči-t-b=ax<sup>w</sup>* *ti?ił sčali? ?ə ti?ił čx<sup>w</sup>əlu?*  
 SCONJ slice-ICS-PASS=NOW DET heart PR DET whale

‘Then Whale’s heart is cut.’

Mink (ES)

What’s going on in this sentence is that the phrase *?ə ti?ił čx<sup>w</sup>əlu?* (“of the whale”) is a periphrastic possessive. Syntactically, it should be part of the constituent headed by *sčali?* (“heart”). Possessor phrases like this appear often without an additional determiner, but I have not yet added any rules or constructions to the grammar intended to handle them. System I finds two parses for this sentence, one due to the error condition caused by the underconstrained SPEC value. The other parse, however, connects the phrase *?ə ti?ił čx<sup>w</sup>əlu?* to the ICS argument of the matrix verb (due to passive marking, the subject should be obliquely marked were it present). This reading, where “whale” cuts “the heart”, is not the correct one in the context of this story. System II finds a corresponding reading without the spurious ambiguity. Now, from my understanding of Lushootseed grammar, this reading looks to be a syntactically valid reading of this sentence.<sup>11</sup> However, the grammar is really failing to find the right parse here since I have no treatment of possessor phrases yet.

I’ve already discussed the main reasons why System I ends up being so much more ambiguous than System II. The ambiguity of System II on this story was in two examples, which fell into one of the classes described above: the underconstraint of the valance-changing lexical rule propagating up and leading to additional, incorrect analyses in sentences of medium length.

---

<sup>11</sup>I believe this is another example of the utility of implemented grammars in finding plausible alternative readings for consideration by grammarians.

| <b>'Mink: original: System I' Coverage Profile</b>  |                    |                    |                      |                          |                      |                         |
|-----------------------------------------------------|--------------------|--------------------|----------------------|--------------------------|----------------------|-------------------------|
| <b>Aggregate</b>                                    | <b>total items</b> | <b>word string</b> | <b>lexical items</b> | <b>distinct analyses</b> | <b>total results</b> | <b>overall coverage</b> |
|                                                     | #                  | $\phi$             | $\phi$               | $\phi$                   | #                    | %                       |
| $10 \leq i\text{-length} < 15$                      | 2                  | 11.00              | 0.00                 | 0.00                     | 0                    | 0.0                     |
| $5 \leq i\text{-length} < 10$                       | 27                 | 6.85               | 8.12                 | 2.50                     | 4                    | 14.8                    |
| $0 \leq i\text{-length} < 5$                        | 23                 | 3.22               | 3.94                 | 1.82                     | 11                   | 47.8                    |
| <b>Total</b>                                        | <b>52</b>          | <b>5.40</b>        | <b>5.28</b>          | <b>2.00</b>              | <b>15</b>            | <b>28.8</b>             |
| <b>'Mink: deduped: System I' Coverage Profile</b>   |                    |                    |                      |                          |                      |                         |
| $10 \leq i\text{-length} < 15$                      | 2                  | 11.00              | 0.00                 | 0.00                     | 0                    | 0.0                     |
| $5 \leq i\text{-length} < 10$                       | 25                 | 6.72               | 8.12                 | 2.50                     | 4                    | 16.0                    |
| $0 \leq i\text{-length} < 5$                        | 21                 | 3.24               | 3.94                 | 1.82                     | 11                   | 52.4                    |
| <b>Total</b>                                        | <b>48</b>          | <b>5.37</b>        | <b>5.28</b>          | <b>2.00</b>              | <b>15</b>            | <b>31.2</b>             |
| <b>'Mink: original: System II' Coverage Profile</b> |                    |                    |                      |                          |                      |                         |
| $10 \leq i\text{-length} < 15$                      | 2                  | 11.00              | 0.00                 | 0.00                     | 0                    | 0.0                     |
| $5 \leq i\text{-length} < 10$                       | 27                 | 6.85               | 8.00                 | 2.50                     | 4                    | 14.8                    |
| $0 \leq i\text{-length} < 5$                        | 23                 | 3.22               | 3.94                 | 1.00                     | 11                   | 47.8                    |
| <b>Total</b>                                        | <b>52</b>          | <b>5.40</b>        | <b>5.12</b>          | <b>1.40</b>              | <b>15</b>            | <b>28.8</b>             |
| <b>'Mink: deduped: System II' Coverage Profile</b>  |                    |                    |                      |                          |                      |                         |
| $10 \leq i\text{-length} < 15$                      | 2                  | 11.00              | 0.00                 | 0.00                     | 0                    | 0.0                     |
| $5 \leq i\text{-length} < 10$                       | 25                 | 6.72               | 8.00                 | 2.50                     | 4                    | 16.0                    |
| $0 \leq i\text{-length} < 5$                        | 21                 | 3.24               | 3.94                 | 1.00                     | 11                   | 52.4                    |
| <b>Total</b>                                        | <b>48</b>          | <b>5.37</b>        | <b>5.12</b>          | <b>1.40</b>              | <b>15</b>            | <b>31.2</b>             |

Table 9.4: The evaluation results for Mink

### 9.2.6 Discussion

Overall, these results highlight the utility of using longer, natural text as testsuites in addition to the target paradigms which are of use in initial development. Case study II (Chapter 8) concluded after presenting two systems and comparing their strengths and weaknesses. At the end of that chapter, System I appeared to be the stronger. Even though the semantics of System I are arguably partially tautologous in certain cases, that system provided coverage of the targeted testsuite with no ambiguity and easily extended to certain data that was problematic (at best) for System II — I refer to the general prefixes attaching to nouns and predicative nouns feeding the subject-relativization rule. However, when looking just a bit more broadly than the targeted testsuite, to the 286 sentences of this evaluation, the advantage of System I fell away because of additional erroneous ambiguity which System II did not suffer from. The constructions which System I handled better than System II were rare or else the grammar overall is still too immature to provide the test (i.e. some other unanalyzed phenomenon co-occurred, blocking the opportunity for System I to parse).

## 9.3 Error Analysis on Unparsed Items

In addition to the error analysis of ambiguity and coverage discussed above, going through the items which failed to parse allowed me to collect a list of reasons for failure, which I have categorized into lexical and syntactic topics. In this section, I go through some of the most important of these (in terms of their blocking of coverage of test items).

### 9.3.1 (Lexical) things not implemented

#### *Morphological marking on nominalized verbs*

Far and away the most common reason for a failure of lexical coverage was due to possessive marking on nominalized verbs. Although I added some rudimentary support for nominalizers that attach to verbs and yield noun-like elements, I did not add support for these nominalized verbs to pass through the possessive lexical rules. In fact, most nominalized verbs in the story

do bear some possessive marker as well (morphology which indicates the nominalized verb's subject).

One interesting note on this topic regards the encoding of the morphologically segmented form which the syntactic grammar accepts as input. Beck considers these nominalizers to be phrase-level clitics and since possessive marking cannot appear on verbs without nominalization, the possessive marking on nominalized verbs must also be a clitic.<sup>12</sup> As such, he distinguishes such phrase-level possessive marking on nominalized verbs from purely lexical possessive marking on nouns in the morpheme-segmented line of the corpus (in the traditional way by segmenting them with a hyphen or an equality sign). This choice interacts in an interesting way with my grammar, because in fact, I currently treat these nominalizers as affixes, handled at the level of lexical analysis by the grammar (they simply have equality signs in their spelling instead of hyphens). So, one analysis which is in principle available for me would be to have the nominalized verbs and underived nouns share the lexical position classes which attach possessive marking. However, I would then need two spellings, one with and one without the equals sign in order to align with the structures in the input data. In fact, in order to approach this phenomenon in a satisfactory way, I would have to decide whether the possessive marking on nouns should have different semantics from the possessive marking on verbs, a topic which, unfortunately, fell out of scope for this dissertation. Instead, I forewent the possessive marking on nominalized verbs and therefore the coverage over the evaluation texts suffered proportionally in consequence.

Related to this failure is the grammar's lack of ability to attach the general prefixes to nominalized verbs. As with possessive marking on verbs, the question of theoretical unity becomes important. This time there are four potential attachment sites: an underived verb, an underived noun, a nominalized verb and a predicative noun — any of the four may bear general prefixes which mark aspectual properties such as habitual, mood-like properties such as subjunctive or even tense-like properties such as past and irrealis. As the case study of Chapter 8 and this evaluation chapter so far have already shown, there is a lot of complexity inherent in the potential

---

<sup>12</sup>Beck (p.c.) also motivates this analysis with examples in which the possessive and nominalizing morphology migrate to S2, appearing on preverbal adverbs instead of on the predicate head.

semantic representations. So, as with possessive marking, I kick the can down the road to the detriment of the current system’s lexical coverage metric.

### *Special grammatical particles*

Lushootseed grammar includes a small class of “predicate particles” (Hess, 1995, 88). These sentence-second elements stand out from most Lushootseed words in their inability to be predicative. Instead, they encode particular grammaticalized functions, such as speaker attitude or the resumptivity of an event. These particles may co-occur with other sentence-second elements, such as the subject markers and the interrogative marker (176).

- (176) *ʔasxəł uʔx<sup>w</sup> čəx<sup>w</sup> ʔu*  
*ʔas-xəł uʔx<sup>w</sup> čəx<sup>w</sup> ʔu*  
 STAT-sick still 2SG.SBJ Q

‘Are you still sick?’

LR1 p. 88

What’s more, the ordering of particles in these multi-particle constructions is not completely fixed but in fact provides subtle effects on the meaning which have not been completely described in primary sources. Four of these particles appear in the evaluation data, so my grammar fails to provide any coverage for such sentences.

### *Numeral classifiers*

Lexical suffixes are also used in Lushootseed to make special forms for numbers when counting specific classes of items. Numerals, when used for counting humans or periods of time, have specific forms which must be used — the former based on one of the reduplication patterns (177), the latter using a particular lexical suffix (178). What’s more, lexical suffixes may also attach to numerals when there is a specific lexical suffix which applies to something being counted (179a) or a more general lexical suffix may apply its semantic class to the numeral (179b) (Beck, n.d.,

351). Since I only implemented the plain series of numerals as adjectives, my grammar fails to provide a lexical analysis for any of these specialized number forms.

- (177) *ʔululub ʔi diiĉu? bəkʷ tiʔiʔ kikəwiĉ*  
 ʔul<ul>ub ʔi di<i>ĉu? bəkʷ tiʔiʔ ki-kəwiĉ  
 ten<hmn> CONJ one:HMN all DET ATTN-hunchback

‘Eleven people and Little Hunchback.’

Basket Ogress (LA)

- (178) a. *ʔixʷ[ət]dat tiʔiʔ [s]dəgʷabaciʔ[s]əxʷ əlgʷə? tiʔiʔ ĉxʷəlu?*  
 ʔixʷ•ət•dat tiʔiʔ s=dəkʷa•abac-il=s=axʷ əlgʷə? tiʔiʔ ĉxʷəlu?  
 three•cls•day DET NMLZ=inside•body-INCH=3.POSS=NOW PL DET whale

‘Three days they are inside Whale.’

Mink (ES)

- b. *buusat tiʔiʔ shuys ʔəsʔista?*  
 buus•at tiʔiʔ s=huyu=s ʔas-ʔista?  
 four•cls DET NMLZ=made=3.POSS STAT-be.like

‘It happened like that four times.’

Basket Ogress (AW)

- (179) a. *cəbagʷiʔəxʷ kʷi ʔudəxʷsaxʷəbabacəds əlgʷə?*  
 cəb•a•gʷiʔ=axʷ kʷi ʔu=dəxʷ=saxʷəb•abac-t=s əlgʷə?  
 two•CONN•canoe=NOW DET HAB=ADNM=jump•body-ICS=3.POSS PL

‘Now what they would jump over are two canoes.’

How daylight was stolen (HM)

- b. *ʔəsʔixwulĉ ti dsʔaĉwu?*  
 ʔas-ʔixw•ulĉ ti d-sʔaĉwu?  
 STAT-three•container DET 1SG.POSS-clam

‘I have three clams.’

Hess (1998)

*Directionals*

Lushootseed features a special system of directionals (Beck, n.d., 151). Before I ran this evaluation, I was unaware that the inflectional markers which combine with the locational preposition (see above §9.1.4) were able to combine with these directionals (180), although I have since found these exact forms discussed in Hess and Hilbert (1995b, 114–115). All such forms went unanalyzed by the grammar in this evaluation.

- (180) *kʷitəxʷ*            *dxʷčəʔkʷ*            *ʔal tiʔəʔ stuləkʷ*  
*kʷitʰ=axʷ*            *dxʷ-čəʔkʷ*            *ʔal tiʔəʔ stuləkʷ*  
 shoreward=NOW CNTRPT-seaward LOC DET river

‘He goes down to the river.’

Bear and Fishhawk (ES)

*Unimplemented morphology*

There are still more grammatical affixes of the Lushootseed grammar which are as yet unimplemented in my grammar. Of these, four appeared regularly in the evaluation data, preventing lexical coverage. Especially notable (in that they represent semantic phenomena relatively unrelated to the valence changing morphology I worked on) were the contained prefix (*dxʷ-*) and the partitive (*ʔil-*). The other two were the other secondary suffix (*-i*) and the middle voice marker (*-b*).

*gʷəʔ phrases*

One aspect of Lushootseed grammar which I didn’t address at all are the *gʷəʔ* phrases, which Beck (n.d., 365) glosses “associative”. These phrases are similar, in some ways, to periphrastic possessives formed with *ʔə* prepositions, which are also unimplemented for this evaluation. I mention these here, in the section of lexical failures because my grammar does contain the preposition *ʔə*. Thus the possessive phrases headed by this preposition do not fail to parse due to lexical coverage, whereas at the time of this evaluation, I have no lexical item *gʷəʔ*.

### *Clausal conjunctions*

Lushootseed grammar has a special form for subject-markers when they appear in a clausal conjunction. The subject-marker in the second of the conjoined clauses, instead of targeting sentence-second position, the specialized subject marker appears interclausally, filling the syntactic role of a conjunction (181). The Grammar Matrix customization system does provide functionality to define coordination strategies for S constituents and this functionality would surely provide a helpful basis for a Lushootseed implementation. However, the fusion with the subject markers would have to be handled specially. Such work was not carried out before this evaluation, so any sentences with these special conjunctive forms failed lexical coverage.

- (181) *day' čəx<sup>w</sup> ʔuʔəʔ' [dx<sup>w</sup>?al] ti?iʔ d?al?al čəda*  
*day' čəx<sup>w</sup> ʔu=?əʔ' dx<sup>w</sup>-?al ti?iʔ d-?al?al čəda*  
 uniquely 2SG.SBJ IRR=COME CNTRPT-at dist 1SG.POSS-house 1SG.CONJ  
*ʔuʔəʔʔubicid*  
*ʔu=?əʔ-tx<sup>w</sup>-bicid*  
 IRR=be.fed-ECS-2SG.OBJ

‘After a while you will come to my house and I will feed you.’

Bear and Fishhawk (ES)

### *Other reduplications*

As mentioned above, reduplication patterns were partially implemented for verbs, but not for nouns. In addition to this, there are reduplication patterns even for verbs which were not implemented. Some of these patterns were not fully analyzed in the corpus (182).<sup>13</sup> In other cases, the morphological pipeline wasn't set up to handle constructions such as double reduplications

---

<sup>13</sup>See §4.7 for details about how reduplicative forms were transformed in preparation for processing by the syntactic grammar.

(183). Any of the reduplication forms outside of single applications of DSTR, ATTN OR DIM.EFF ON verbs caused failures in lexical coverage during the evaluation.

- (182) *ʔuʔəxičədəx<sup>w</sup>*                      *čəx<sup>w</sup>*    *sbiaw*  
 ʔu-ʔəxi-čəd=ax<sup>w</sup>                      čəx<sup>w</sup>    sbiaw  
 PFV-RDP-what.happen=now 2SG.SBJ coyote

‘What are you doing, Coyote?’

Coyote and Big Rock (ES)

- (183) *g<sup>w</sup>a·dadg<sup>w</sup>adəx<sup>w</sup>*                      *əlg<sup>w</sup>əʔ*  
 g<sup>w</sup>ad-ad-g<sup>w</sup>ad=ax<sup>w</sup>                      əlg<sup>w</sup>əʔ  
 DSTR-DIM.EFF-talk=now PL

‘They chat a bit.’

Bear and Fishhawk (ES)

### 9.3.2 Imperative

Lushootseed has a second-person subject pronoun which indicates imperative sentential force (184). Imperatives are relatively rare in the corpus and presumably follow the same syntactic patterns as the other main clause subject markers (requiring sentence-second position), but due to the lack of examples with adverbials which would test the matter, this hypothesis is not fully established (Beck, n.d., 413). Nevertheless, a sentence-second implementation will be easy for my grammar system, given my work on the subject markers described in Chapter 6. This item was not implemented simply because I had not run across it in my studies until this evaluation had begun.

- (184) *dəg<sup>w</sup>ag<sup>w</sup>il*    *ʔi*                      *dsuq<sup>w</sup>suq<sup>w</sup>aʔ*  
 dək<sup>w</sup>a-ag<sup>w</sup>il    ʔi                      d-suq<sup>w</sup>-suq<sup>w</sup>aʔ  
 inside-AUTO 2PL.IMP 1SG.POSS-DSTR-younger.sibling

‘Get inside my younger brothers!’

Coyote and Big Rock (ES)

### 9.3.3 Parse errors

The other set of errors which I reviewed after carrying out the evaluation were those where there was lexical coverage for each word-token in the input, but the syntactic rules of the grammar failed to find a syntactic analysis. I have identified four types of parse errors which implicated in most of these parse failures.

#### *Initial determiners*

My grammar operates on a strict requirement that on one hand, all noun phrases in argument position require a determiner while nouns which appear in predicate position of matrix clauses do not allow a determiner. While this is the most common pattern in the data I have observed, the facts of Lushootseed do not bear this out generally. Several of the evaluation sentences begin with determiner-marked nouns (185). My initial reaction to this data is that it evinces a need for a left-dislocation construction which would be complementary to the predicative noun rule. Additionally, there are speakers who consistently produce noun phrases that lack determiners (Zahir 2019, p.c.), although such examples do not appear in this evaluation data.

(185) a. *tiʔəʔ sčətxʷəd gʷəl ʃʷuʹ lʷiʔibibəš*

*tiʔəʔ sčətxʷəd gʷəl ʃʷuʹ lʷiʔib-ʔibəš*

DET bear        CONJ only    HAB=DSTR-travel

‘This Black Bear, he would just wander around.’

Black Bear and Ant (ES)

b. *tiʔiʔ sbiaw gʷəl ʔuʃʷəxʷ*

*tiʔiʔ sbiaw gʷəl ʔuʃʷ=əxʷ*

DET coyote    CONJ go=now

‘That Coyote, he goes.’

Coyote and Big Rock (ES)

9.3.4 *Appositive possessors*

In “Bear and FishHawk” and in “Mink”, there are several examples of an appositive-like construction where one of the NPs is a possessive construction. The storyteller continually reminds us that Bear and FishHawk are mutual friends by referring to either as *sya?ya?s* (“his friend”) (186a –186c). Mink’s cousin is also usually introduced with a possessive *su?suqʷa?s* (“his younger cousin”) which reminds us of his relationship to the protagonist.

- (186) a. *ʔastagʷaxʷaxʷ ti?iʔ sya?ya?s sčətxʷəd*  
*ʔas-tagʷaxʷ=axʷ ti?iʔ sya?ya?-s sčətxʷəd*  
 STAT-hungry=now DET friend-3.POSS bear

‘His friend Black Bear is hungry.’

- b. *gʷəl ʔčisəb ʔə ti?iʔ ʔayəds čixčix*  
*gʷəl ʔčil-s-b ʔə ti?iʔ ʔayəd-s čixčix*  
 SCONJ arrive-ALTV-PASS PR DET companion-3.POSS fish.hawk

‘And he is arrived at by his friend Fish Hawk.’

- c. *huy cuucəxʷ ti?iʔ sya?ya?s čixčix*  
*huy cut-c=axʷ ti?iʔ sya?ya?-s čixčix*  
 SCONJ say-ALTV=now DET friend-3.POSS fish.hawk

‘Then, he says to his friend, Fish Hawk,’

Bear and FishHawk (ES)

- d. *+ti?iʔ bibščəb ʔi ti?iʔ su?suqʷa?s tətyika*  
*ti?iʔ bi-bəščəb ʔi ti?iʔ su?-suqʷa?-s tətyika*  
 DET ATTN-mink CONJ DET ATTN-younger.sibling-3po Tutyika

‘...Little Mink and his little younger cousin, Tutyika.’

Mink (ES)

This construction was common enough in these stories to hamper coverage considerably (being that the current grammar has no analysis for it). However, it does not appear in either of

the other two stories in this evaluation, so I make no claim that this is a common construction in general.

*Performative sentences*

In addition to the refrains of songs (mentioned above), there are some other aspects of Lushootseed storytelling style which are not completely compatible with the default root-conditions of a Matrix-derived grammar. On these I noted especially performative type sentences. For example, the incantation of Fishhawk which is appears several times in the story is evocative of its effect in its repetition (187), but it's unclear to me how to squeeze this into a typical grammatical constituent in a way which would generally be useful for testing the grammar.<sup>14</sup>

(187) *ciχəb ciχəb ciχəb ciχəb ciχəb*  
*ciχəb ciχəb ciχəb ciχəb ciχəb*  
 drip drip drip drip drip

“Drip, drip, drip, drip, drip.”

Bear and Fishhawk (ES)

Similarly, in “Bear and Ant,” Mr. Sam repeats the phrase *ləcuyayus* multiply (188). For sentences such as these which bear stylistic and rhetorical features whose analysis is not usually included in the domain of sentential syntax, I was usually able to get lexical coverage, but failed to find a spanning parse.

(188) *ləcuyayus ləcuyayus ləcuyayus ckʷaqid ləcuyayus*  
*ləcu-yayus ləcu-yayus ləcu-yayus ckʷaqid ləcu-yayus*  
 CONT-work CONT-work CONT-work always CONT-work

‘She is working, working, working, always working.’

Bear and Ant (ES)

---

<sup>14</sup>There are, of course, ways to cover such utterances. A complete treatment would take me into an analysis of Lushootseed discourse which lies out of scope for my this project.

*Sentence segmentation and corpus lines*

Upon reading the texts, I have some doubts as to whether the segmentation into lines in the corpus I am working from is reflective of a traditional notion of sentence division or reflective of the phonological, or intonational, structure which the storyteller uses (where the two may not coincide). Lines 6–7 of “Bear and Ant” are illustrative of certain aspects of the matter (189). I have three observations regarding this pair of lines. First, the original publication of these texts, the unglossed version published in LR1 (Hess, 1995, 143), bears additional punctuation which has been removed from Beck’s annotated corpus: commas separate off sentential conjunctions, colons introduce quotations and most lines end in a full stop, including both (189a) and (189b). Second, these texts were carefully edited by Hess and Sam after the recording, in some cases material which is not heard on the recording has been inserted (Beck, n.d.), line segmentation was surely chosen with care. Nevertheless, thirdly, Beck’s free translation of (189a) ends with a comma, beginning an English sentence which ends with the free translation of (189b). I think that this provides an oblique suggestion that the sentences are more closely connected than the line segmentation suggests.

- (189) a. *łiqag<sup>w</sup>ilax<sup>w</sup>      ti?it<sup>h</sup> sčətx<sup>w</sup>əd tul?al      ti?it<sup>h</sup> ?al?als*  
*łiqi-ag<sup>w</sup>il=ax<sup>w</sup>      ti?it<sup>h</sup> sčətx<sup>w</sup>əd tul’-?al      ti?it<sup>h</sup> ?al?al-s*  
 emerge-AUTO=NOW DET bear      CNTRFG-LOC DET house-3.POSS

‘Black Bear comes out of his house,’

- b. *ti?it<sup>h</sup> ?al?als      łudax<sup>w</sup>?ux<sup>w</sup>s      ?al ti?it<sup>h</sup> pədtəs*  
*ti?it<sup>h</sup> ?al?al-s      łu=dax<sup>w</sup>=?ux<sup>w</sup>=s      ?al ti?it<sup>h</sup> pədtəs*  
 DET house-3.POSS HAB=ADNM=GO=3.POSS LOC DET winter

‘his house where he would go in the winter.’

Bear and Ant (ES), lines 6 – 7

If the latter item in (189) were an appositive of the final noun phrase in the former, my grammar would come closer to parsing it (after adding rules for NP apposition) since my grammar

currently requires determiners on all NPs. On the other hand, the lack of determiner between *ʔalʔals* and *ʔudax<sup>w</sup>ʔux<sup>w</sup>s*, to me, rekindles the idea of treating this as a left-dislocation construction (as mentioned above in the section on initial-determiners). But this pair isn't the only example which lead me to question the alignment of lines with sentences. There is a line in "Mink" which contains only the word *ʔu* ("okay"), a direct quote introduced by the previous line (190). In Hess' original, in this case, the first line ends with a comma, the latter with a full stop.

- (190) a. *huy cutax<sup>w</sup> tiʔiʔ bibščəb*  
           *huy cut=ax<sup>w</sup> tiʔiʔ bi-bəščəb*  
           SCONJ say=now DET ATTN-mink

‘Then Little Mink says,’

- b. *ʔu*  
       *ʔu*  
       INTJ

‘Okay.’

Mink (ES)

Another example comes from lines which have the sentential connective *g<sup>w</sup>əl* in the middle, ostensibly joining two stand-alone clauses. The two examples above in (185) have this property. Generally in the corpus, *g<sup>w</sup>əl* is glossed as a sentential connective, and was implemented for the evaluation as a sentence-initial adverb. If the main refrain of Ant's song *ʔaxil g<sup>w</sup>əl bəʔaxil* (159) were presented to my grammar as two separate items *ʔaxil* ("Night falls") and *g<sup>w</sup>əl bəʔaxil* ("and it becomes day again"), it would have provided a sensible parse for each. However, since the current system has no analysis for sentential connectives mid-sentence, I wasn't able to provide any coverage for such items.

The main takeaway from this section is not that I am arguing strongly for a different line segmentation than the one in the corpus, but more so that I believe there may be a misalignment between the expectations which I have encoded into the grammar structures and those which came to bear on producing the line segmentation. I think that it boils down to a matter which I

need to study further in order to decide how to properly configure the grammar for analysis of these texts.

#### ***9.4 Conclusion***

The implemented grammar system I have been presenting in this dissertation was initially constructed and evaluated using small, targeted testsuites which attempt to isolate analytical dimensions and list all the variants of a paradigm. These testsuites include negative examples where I was able to reasonably construct them from my reading of Lushootseed Reader volume I and other materials after presenting them to language experts for their validation. This evaluation has expanded the scope and type of data by using the grammar-system to parse naturally occurring texts which are transcriptions of traditional stories. In addition to the evaluation of the system as a whole, I have used this exercise to compare the two approaches to noun-phrases in predicative contexts which were introduced in Chapter 8.

My opinion is that overall the system performed well. Given the limited implementation, the coverage values were usually greater than 50% for shorter sentences and usually greater than 25% on medium length sentences while maintaining low ambiguity (modulo the caveats presented in the discussion above). The evaluation also highlighted a list of unimplemented language features which provide direction for future development.

The comparison of System I and System II validated the importance of the ability of System I to treat general prefixes on nouns in embedded contexts. It also highlighted how an analytical oversight which wasn't apparent in the smaller, targeted testsuite can lead to large amounts of spurious ambiguity in a more expansive test. This unconstrained SPEC value in System I is something which will have to be addressed if it's going to remain a viable option for the grammar system's eventual direction. Overall, the evaluation provided both concrete next-steps and a baseline against which future development can be benchmarked.

## Chapter 10

### CONCLUSION

In this document I have presented the broad outlines of a system for generating and iteratively improving a computational grammar for Lushootseed. I have provided a rudimentary approach to morphophonology and morphosyntax and validated the analyses on Lushootseed testsuites. In most cases, the specifics of the linguistic analysis itself was due to preceding scholars, my contribution was the fleshing out of those analyses into fully-fledged, mathematical structures and rules which are composable into instructions for a computing machine.

My system relies on theories, tools and techniques which have been developed, for the most part, in the context of the analysis of non-American languages. This dissertation provides detailed case-studies exhibiting the use of this framework on the phenomenon of Lushootseed valence-changing affixes and on the issues surrounding predicative nouns and non-predicative verbs. In the former case, treating the valence increasing affixes as derivational morphology (as opposed to inflectional) and adhering to a model of sub-atomic combinatorics for the creation of elementary predications led to a refinement of a particular technical principle of MRS construction. In the latter case, I explore two approaches to a concrete implementation of the fact that Lushootseed (like other Salish languages) allows great syntactic flexibility for nouns to appear in predicative positions and verbs in argument positions. I explore the interaction of these two approaches with morphological markers which contribute meaning typically associated with event-like semantics but which can appear on either nouns or verbs. For the most part, the theory and framework held up well and the implementation highlighted analytical options which presented particular tradeoffs for the grammar writer more so than any deep inadequacies of the underlying theory.

This document also includes an evaluation chapter which puts the syntactic implementation to the task of analyzing real text (as opposed to the paradigmatic examples of the test suites).

Because neither of the implementations I explored in Chapter 8 was completely satisfactory, I carried both of these over into the evaluation, further contrasting the differences between the analyses. The contrast was overshadowed to some degree by errors in the implementation which were also revealed by the evaluation exercise. Much of Chapter 9 is dedicated to error analysis which outlines the next steps for this project. The results of the evaluation task also implicitly create an importance ranking for these next steps, since they highlighted the relative frequency of unimplemented phenomena.

To summarize everything up at a high level, I return to the anecdote I mentioned in the introduction. This implementation represents a first attempt at a basket, one which, in many ways, simply doesn't hold water. Nevertheless, it also represents a base upon which iterative improvement can be made. In this way, I hope it is useful to others who may come across it. To that end, one further loose end should be addressed, that of the availability of the grammar code and software I have developed. I hope that in the future I will be able to completely release this grammar project in a way which maximizes its utility for others to study and improve upon. However, I also must recognize that Lushootseed does not belong to me but to the first people of Puget Sound. For that reason, my next step in making the grammar available is to coordinate with the intertribal organization which Vi Hilbert founded in order to coordinate and disseminate materials on Lushootseed Language studies, the non-profit organization Lushootseed Research.<sup>1</sup>

---

<sup>1</sup><http://lushootseedresearch.org>

## BIBLIOGRAPHY

- Abeillé, A. and Godard, D. (2001). A class of “lite” adverbs in french. *Amsterdam Studies in the Theory and History of Linguistic Science Series 4*, pages 9–26.
- Bach, E. (1976). An extension of classical transformational grammar. *Mimeo*.
- Bagley, C. B. (1915). Journal of occurrences at Nisqually House, 1833. *The Washington Historical Quarterly*, 6:179–197.
- Ballard, A. C. (1929). Mythology of Southern Puget Sound: legends shared by tribal elders. *University of Washington Publications in Anthropology*, 3(2):31–150.
- Barthmaier, P. (1998). Lushootseed vowels: a primary phonetic study. In *Proceedings of the 33rd Annual International Conference on Salish and Neighboring Languages (ICSNL)*. University of British Columbia.
- Barwise, J. and Cooper, R. (1981). Generalized quantifiers and natural language. *Linguistics and Philosophy*, 4:159–219.
- Barwise, J. and Perry, J. (1983). *Situations and Attitudes*. The MIT Press, Cambridge, MA.
- Bates, D. (1999). Distance in narrative time and space: Aspect markers and determiner choice in Martha Lamont’s “pheasant and raven”. In *Proceedings of the 34th Annual International Conference on Salish and Neighboring Languages (ICSNL)*. University of British Columbia.
- Bates, D. (2002). Narrative functions of past-tense marking in a Lushootseed text. In *Proceedings of the 37th Annual International Conference on Salish and Neighboring Languages (ICSNL)*, volume 9 of *University of British Columbia Working Papers in Linguistics*.
- Bates, D. and Hess, T. (2001). Tense or aspect? a prefix of future time in Lushootseed. In *Proceedings of the 36th Annual International Conference on Salish and Neighboring Languages (ICSNL)*, volume 6 of *University of British Columbia Working Papers in Linguistics*, pages 26–36.
- Bates, D., Hess, T., and Hilbert, V. (1994). *Lushootseed Dictionary*. University of Washington Press, Seattle.

- Bates, D. and Lonsdale, D. (2010). Recovering and updating legacy dictionary data. In *Proceedings of the 44th Annual International Conference on Salish and Neighboring Languages (ICSNL)*, volume 27 of *University of British Columbia Working Papers in Linguistics*, pages 1–12.
- Beck, D. (1999). Words and prosodic phrasing in Lushootseed narrative. In Hall, T. A. and Kleinhenz, U., editors, *Studies on the Phonological Word*, pages 23–46. Benjamins, Amsterdam.
- Beck, D. (2009). A taxonomy and typology of Lushootseed valency-increasing suffixes. *International Journal of American Linguistics*, 75(4):533–569.
- Beck, D. (2013). Uni-directional flexibility and the noun–verb distinction in Lushootseed. In Rijkhoff, J. and van Lier, E., editors, *Flexible word classes: A typological study of underspecified parts-of-speech*, pages 185–220. Oxford University Press, Oxford.
- Beck, D. (n.d.). Lushootseed morphosyntax. University of Alberta.
- Beck, D. and Hess, T. (2014). *Tellings from Our Elders: Lushootseed syəyəhub. Volume 1, Snohomish texts*. UBC Press, Vancouver.
- Beck, D. and Hess, T. (2015). *Tellings from Our Elders: Lushootseed syəyəhub. Volume 2. Tales from the Skagit Valley*. UBC Press, Vancouver.
- Beesley, K. R. and Karttunen, L. (2003). *Finite State Morphology*. CSLI Publications, Stanford.
- Bender, E. M. (2008). Grammar engineering for linguistic hypothesis testing. In Gaylord, N., Palmer, A., and Ponvert, E., editors, *Proceedings of the Texas Linguistics Society X Conference: Computational Linguistics for Less Studied Languages*, pages 16–36. Stanford CSLI Publications ONLINE.
- Bender, E. M. (2010). Reweaving a grammar for Wambaya: A case study in grammar engineering for linguistic hypothesis testing. *Linguistic Issues in Language Technology*, 3:1–34.
- Bender, E. M., Drellishak, S., Fokkens, A., Goodman, M. W., Mills, D. P., Poulson, L., and Saleem, S. (2010). Grammar prototyping and testing with the LinGO Grammar Matrix Customization System. In *Proceedings of the ACL 2010 System Demonstrations*, pages 1–6, Uppsala, Sweden. Association for Computational Linguistics.

Bender, E. M. and Flickinger, D. (2005). Rapid prototyping of scalable grammars: Towards modularity in extensions to a language-independent core. In *Proceedings of the 2nd International Joint Conference on Natural Language Processing IJCNLP-05 (Posters/Demos)*, Jeju Island, Korea.

Bender, E. M., Flickinger, D., and Oepen, S. (2002). The Grammar Matrix: An open-source starter-kit for the rapid development of cross-linguistically consistent broad-coverage precision grammars. In Carroll, J., Oostdijk, N., and Sutcliffe, R., editors, *Proceedings of the Workshop on Grammar Engineering and Evaluation at the 19th International Conference on Computational Linguistics*, pages 8–14, Taipei, Taiwan.

Bender, E. M., Flickinger, D., and Oepen, S. (2011). Grammar engineering and linguistic hypothesis testing: Computational support for complexity in syntactic analysis. In Bender, E. M. and Arnold, J. E., editors, *Language from a Cognitive Perspective: Grammar, Usage and Processing*, pages 5–29. CSLI Publications, Stanford.

Bender, E. M. and Good, J. (2005). Implementation for Discovery: A bipartite lexicon to support morphological and syntactic analysis. *Chicago Linguistic Society*, 41(The Panels).

Bender, E. M. and Good, J. (2010). A Grand Challenge for Linguistics: Scaling up and integrating models. Technical Report SBE 2020 Id: 81, National Science Foundation.

Bender, E. M., Goodman, M. W., Crowgey, J., and Xia, F. (2013). Towards creating precision grammars from interlinear glossed text: Inferring large-scale typological properties. In *Proceedings of the 7th Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities*, pages 74–83, Sofia, Bulgaria. Association for Computational Linguistics.

Bender, E. M. and Langendoen, D. T. (2010). Computational Linguistics in Support of Linguistic Theory. *Linguistic Issues in Language Technology*, 3(2):1–31.

Bender, E. M., Poulson, L., Drellishak, S., and Evans, C. (2007). Validation and regression testing for a cross-linguistic grammar resource. In *ACL 2007 Workshop on Deep Linguistic Processing*, pages 136–143, Prague, Czech Republic. Association for Computational Linguistics.

Bevans, C. I., editor (1968). *Treaties and Other International Agreements of the United States of America, 1776–1949*. United States Department of State.

Bickel, B., Comrie, B., and Haspelmath, M. (2008). The Leipzig glossing rules. conventions for interlinear morpheme glosses. <http://www.eva.mpg.de/lingua/resources/glossing-rules.php>. Revised 2015, Accessed: 2019-08-24T18:51+00:00.

- Bierwert, C. (1996). *Lushootseed Texts: an introduction to Puget Salish Narrative Aesthetics*. University of Nebraska Press.
- Bird, S. and Simons, G. (2003). Seven dimensions of portability for language documentation. *Language*, pages 557–582.
- Bloomfield, L. and Bolling, G. M. (1927). What symbols shall we use? *Language*, 3:123–129.
- Boas, F. and Haeblerlin, H. (1927). Sound shifts in Salishan dialects. *International Journal of American Linguistics*, 4(2):117–136.
- Bond, F., Oepen, S., Nichols, E., Flickinger, D., Velldal, E., and Haugereid, P. (2011). Deep open-source machine translation. *Machine Translation*, 25:87–105.
- Bos, J. (1995). Predicate logic unplugged. In *Proceedings of the 10th Amsterdam Colloquium*, pages 133–142, Amsterdam.
- Bossong, G. (1992). Reflections on the history of the study of universals: the example of the *partes orationis*. In Kefer, M. and van der Auwera, J., editors, *Meaning and grammar: cross-linguistic perspectives*, pages 3–16. Mouton de Gruyter, Berlin.
- Bray, T., Paoli, J., Sperberg-McQueen, C. M., Maler, E., and Yergeau, F. (2008). <http://www.w3.org/TR/2008/REC-xml-20081126/>. Accessed 2019-05-05T17:52+00:00.
- Bresnan, J. and Mchombo, S. A. (1995). The Lexical Integrity Principle: Evidence from Bantu. *Natural Language & Linguistic Theory*, 13(2):181–254.
- Bruening, B. (2018). The lexicalist hypothesis: both wrong and superfluous. *Language*, 94(1):1–42.
- Burton, S., Davis, H., Jacobs, P., Tamburri Watt, L., and Wiltschko, M. (2001). A boy, a dog and a frog. *International Conference on Salish and Neighboring Languages*, 36:37–54.
- Carlson, R. L. (1990). History of research in archeology. In Suttles, W., editor, *Handbook of North American Indians Volume 7, Northwest Coast*, volume 7 of *Handbook of North American Indians*, pages 107–115. Smithsonian Institution.
- Chalcraft, E. L. (2004). *Assimilation's Agent: my life as a superintendent in the Indian Boarding School System*. University of Nebraska Press, Lincoln and London. Edited and with an introduction by Cary C. Collins.

- Chierchia, G. (1989). Structured meanings. In Chierchia, G., Partee, B. H., and Turner, R., editors, *Properties, Types and Meaning, Volume II: Semantic Issues*, volume 39 of *Studies in Linguistics and Philosophy*. Kluwer Academic Publishers, Dordrecht.
- Chirouse, E. C. (1880). A short method to learn the Snohomish Indian language in 14 lessons. Manuscript No. HHB [P-B86] in the Bancroft Library of University of California, Berkely.
- Chomsky, N. (1957). *Syntactic Structures*. Mouton, The Hague/Paris.
- Chomsky, N. (1959). On certain formal properties of grammars. *Information and Control*, 2(2):137–167.
- Chomsky, N. (1963). Formal properties of grammars. In Luce, D. R., Bush, R. R., and Galanter, E., editors, *Handbook of Mathematical Psychology*, pages 323–418. John Wiley and Sons, New York and London.
- Chomsky, N. (1965). *Aspects of the Theory of Syntax*. The MIT Press, Cambridge, Massachusetts.
- Chomsky, N. (1966). *Cartesian Linguistics: A Chapter in the History of Rationalist Thought*. Harper and Row, New York.
- Chomsky, N. (1981). *Lectures on Government and Binding*. Foris Publications, Dordrecht.
- Chomsky, N. and Halle, M. (1968). *The Sound Pattern of English*. Studies in Language. Harper & Row, New York.
- City of Seattle (2017). Seattle city webpage about Mee-Kwa-Mooks park. <https://www.seattle.gov/parks/find/parks/me-kwa-mooks-park>. Accessed: 2017-11-23T00:00+00:00.
- Clinton, R. (2007). Code of Indian offenses. <http://robert-clinton.com/wordpress/?p=72>. Accessed: 2018-06-05T22:51:06+00:00.
- Copestake, A. (1992). *The Representation of Lexical Semantic Information*. PhD thesis, University of Sussex.
- Copestake, A. (2002). *Implementing Typed Feature Structure Grammars*. Cambridge University Press.
- Copestake, A. (2006). Robust Minimal Recursion Semantics. Unpublished Manuscript.

Copestake, A. (2007a). Applying robust semantics. In *10th Conference of the Pacific Association for Computational Linguistics*, pages 1–12.

Copestake, A. (2007b). Semantic composition with (Robust) Minimal Recursion Semantics. In *ACL 2007 Workshop on Deep Linguistic Processing*, pages 73–80. Association for Computational Linguistics.

Copestake, A. (2009). Slacker semantics: Why superficiality, dependency and avoidance of commitment can be the right way to go. In *Proceedings of the 12th Conference of the European Chapter of the ACL (EACL 2009)*, pages 1–9. Association for Computational Linguistics.

Copestake, A., Emerson, G., Goodman, M. W., Horvat, M., Kuhnle, A., and Muszyńska, E. (2016). Resources for building applications with dependency Minimal Recursion Semantics. In *LREC 2016 Proceedings*, pages 1240–1247.

Copestake, A., Flickinger, D., Pollard, C., and Sag, I. A. (2005). Minimal Recursion Semantics: An introduction. *Research on Language & Computation*, 3(4):281–332.

Copestake, A., Lascarides, A., and Flickinger, D. (2001). An algebra for semantic construction in constraint-based grammars. In *Proceedings of the 39th Annual Meeting of the Association for Computational Linguistics*, Toulouse, France.

Craig, C. G., editor (1984). *Noun classes and categorization*. Typological Studies in Language. John Benjamins, Amsterdam.

Crowgey, J. and Bender, E. M. (2011). Analyzing interacting phenomena: Word order and negation in Basque. In Müller, S., editor, *Proceedings of the HPSG11 conference*. Stanford: CSLI Publications.

Crysmann, B. (2010). Discontinuous negation in Hausa. In Müller, S., editor, *The Proceedings of the 17th International Conference on Head-Driven Phrase Structure Grammar*, pages 269–287. Stanford: CSLI Publications.

Czaykowska-Higgins, E. (1996). *What's in a word? : word structure in Moses-Columbia Salish (Nxa'amxcin)*. Voices of Rupert's Land, Winnipeg.

Czaykowska-Higgins, E. and Kinkade, D. M. (1998). Salish languages and linguistics. In Czaykowska-Higgins, E. and Kinkade, D. M., editors, *Salish Languages and Linguistics*, pages 1–71. Mouton de Gruyter, New York.

- David, P. (1985). Clio and the economics of QWERTY. *American Economic Review*, 75(2):332–337.
- Davidson, D. (1967). The logical form of action sentences. *The logic of decision and action*, pages 81–95.
- Davidson, D. (1985). Adverbs of action. In Vermazen, B. and Merrill, H., editors, *Essays on Davidson: Actions and Events*. Clarendon, Oxford.
- Davis, H. and Matthewson, L. (2009). Issues in Salish syntax and semantics. *Language and Linguistics Compass*, 3(4):1097–1166.
- Dover, H. S. (2013). *Tulalip From My Heart*. University of Washington Press, Seattle and London.
- Dowty, D. (1991). Thematic proto-roles and argument selection. *Language*, 67(3):547–619.
- Dowty, D. R. (1989). On the semantic content of the notion of ‘thematic role’. In Chierchia, G., Partee, B. H., and Turner, R., editors, *Properties, Types and Meaning, Volume II: Semantic Issues*, volume 39 of *Studies in Linguistics and Philosophy*. Kluwer Academic Publishers, Dordrecht.
- Drellishak, S. and Bender, E. M. (2005). A coordination module for a crosslinguistic grammar resource. In Müller, S., editor, *Proceedings of the 12th International Conference on Head-Driven Phrase Structure Grammar*, Stanford. CSLI.
- Eco, U. (1993). *The search for the perfect language*. Blackwell, Oxford.
- Fillmore, C. J. (1968). The case for case. In Bach, E. and Harms, R. T., editors, *Universals in Linguistic Theory*, pages 1–88. Holt, Rinehart, and Winston, New York.
- Flickinger, D. (2000). On building a more efficient grammar by exploiting types. *Natural Language Engineering*, 6(1):15–28.
- Flickinger, D. and Bender, E. M. (2003). Compositional semantics in a multilingual grammar resource. In Bender, E. M., Flickinger, D., Fouvry, F., and Siegel, M., editors, *Proceedings of the Workshop on Ideas and Strategies for Multilingual Grammar Development, ESSLLI 2003*, pages 33–42, Vienna, Austria.
- Flickinger, D., Bender, E. M., and Oepen, S. (2003). MRS in the LinGO Grammar Matrix: A practical user’s guide. accessed: 2018-10-15T05:38:56+0000.

Flickinger, D., Bender, E. M., and Oepen, S. (2014a). ERG semantic documentation. <http://www.delph-in.net/esd>. Accessed on 2018-10-14.

Flickinger, D., Bender, E. M., and Oepen, S. (2014b). Towards an encyclopedia of compositional semantics: Documenting the interface of the English Resource Grammar. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC-2014)*, pages 875–881. European Language Resources Association (ELRA).

Flickinger, D., Bender, E. M., and Packard, W. (2016). English resource semantics. In *Tutorial Abstracts, NAACL HLT 2016, The 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, San Diego California, USA, June 12-17, 2016*, pages 1–5.

Fokkens, A. (2010). Documentation for the Grammar Matrix word order library. Technical report, Saarland University.

Fokkens, A. (2011). Metagrammar engineering: Towards systematic exploration of implemented grammars. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, pages 1066–1076, Portland, Oregon, USA. Association for Computational Linguistics.

Fokkens, A. (2014). *Enhancing Empirical Research for Linguistically Motivated Precision Grammars*. PhD thesis, Universität des Saarlandes, Saarbrücken.

Fokkens, A., Avgustinova, T., and Zhang, Y. (2012). CLIMB grammars: three projects using metagrammar engineering. In *Proceedings of the 8th LREC*, Istanbul, Turkey.

Fokkens, A., van Erp, M., Postma, M., Pedersen, T., Vossen, P., and Freire, N. (2013). Offspring from reproduction problems: What replication failure teaches us. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*, Sofia, Bulgaria.

Fortescue, M. (2009). Lexical suffixes and the position of Proto-Wakashan within the north-west coast linguistic area. *The Canadian Journal of Linguistics*, 54:27–41.

Francez, N. and Wintner, S. (2012). *Unification grammars*. Cambridge University Press, New York, NY.

Garson, J. (2018). Modal logic. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, fall 2018 edition.

Gates, Charles, M., editor (1940). *Messages of the Governors of the Territory of Washington to the Legislative Assembly, 1854–1889*, volume 12 of *University of Washington Publications in the Social Sciences*. University of Washington Press, Seattle.

- Gazdar, G. (1981). Unbounded dependencies and coordinate structure. In Savitch, W., Bach, E., W., M., and Safran-Naveh, G., editors, *The Formal Complexity of Natural Language*, volume 33 of *Studies in Linguistics and Philosophy (formerly Synthese Language Library)*. Springer, Dordrecht.
- Gazdar, G. (1983). Phrase structure grammars and natural languages. In *Proceedings of the Eighth international joint conference on Artificial intelligence*, volume 1.
- Geach, P. T. (1962). *Reference and Generality: an examination of some medieval and modern theories*. Cornell University Press, Ithaca and London.
- Gerds, D. (1988). *Object and Absolutive in Halkomelem Salish*, volume 55 of *Routledge Library Editions: Linguistics*. Garland Publishing, Inc., New York.
- Gerds, D. (1999). The combinatory properties of Halkomelem lexical suffixes. *Annual Meeting of the Berkeley Linguistics Society*, 25(1):337–347.
- Gerds, D. (2003). Halkomelem lexical suffixes. *International Journal of American Linguistics*, 69(4):345–356.
- Gerds, D. and Hukari, T. (2008). The expression of noun phrases in Halkomelem texts. *Anthropological Linguistics*, 50(3/4):324–364.
- Gibbs, G. (1863a). *Alphabetical vocabularies of the Clallam and Lummi*. Cramoisy Press, New York.
- Gibbs, G. (1863b). *Dictionary of the Chinook Jargon or The Trade Language of Oregon*. Smithsonian Institution, Washington.
- Gibbs, G. (1877). Dictionary of the Niskwalli. In Powell, J. W., editor, *Contributions to North American Ethnology*, volume 1, pages 286–361. Government Printing Office, Washington.
- Goodman, M. W. (2013). Generation of machine-readable morphological rules from human-readable input. In Song, S. and Crowgey, J., editors, *UW Working Papers in Linguistics*, volume 30.
- Gruber, J. S. (1965). *Studies in Lexical Relations*. PhD thesis, Massachusetts Institute of Technology.
- Gunther, E. (1945). Ethnobotany of Western Washington. *University of Washington Publications in Anthropology*, 10(1).

- Haeberlin, H. and Gunther, E. (1930). The Indians of Puget Sound. *University of Washington Publications in Anthropology*, 4(1):1–84.
- Hale, K. (1992). Endangered languages: On endangered languages and the safeguarding of diversity. *Language*, 68:1–42.
- Haspelmath, M. (2012). How to compare major word-classes across the world’s languages. In *UCLA Working Papers in Linguistics: Theories of Everything*, volume 17, pages 109–130.
- Herbelot, A. (2013). What is in a text, what isn’t, and what this has to do with lexical semantics. In *Proceedings of the 10th International Conference on Computational Semantics (IWCS 2013)*, pages 321–327.
- Herbelot, A. and Copestake, A. (2006). Acquiring ontological relationships from Wikipedia using RMRS. In *Proceedings of the ISWC 2006 Workshop on Web Content*.
- Hess, T. (1973). Agent in a Coast Salish language. *International Journal of American Linguistics*, 39:89–94.
- Hess, T. (1976). *Dictionary of Puget Salish*. University of Washington Press, Seattle.
- Hess, T. (1977). Lushootseed dialects. *Anthropological Linguistics*, 19(9):403–419.
- Hess, T. (1984). Morphophonemic spelling for pedagogical and other practical purposes. *Working Papers of the Linguistics Circle of the University of Victoria*, 4(2).
- Hess, T. (1992). ʔəsʔistə. In Yoder, J., editor, *dxʷʔal taqʷšəblu tulʔal ti syəyaʔyaʔs: Writings about Vi Hilbert by her friends*, pages 25–27. Lushootseed Press.
- Hess, T. (1993). A schema for the presentation of Lushootseed verb stems. In Mattina, A. and Timothy, M., editors, *American Indian linguistics and ethnography in honor of Lawrence C. Thompson*, number 10 in University of Montana Occasional Papers in Linguistics, pages 113–125. University of Montana, Missoula, MT.
- Hess, T. (1995). *Lushootseed Reader with Introductory Grammar: vol. I: Four Stories from Edwards Sam*. Number 11 in University of Montana Occasional Papers in Linguistics. University of Montana, Missoula.
- Hess, T. (1998). *Lushootseed Reader with Intermediate Grammar: vol. II: Four Stories from Martha Lamont*. Number 14 in University of Montana Occasional Papers in Linguistics. University of Montana, Missoula.

- Hess, T. and Bates, D. (1998). Semantic role assignment in Lushootseed causatives. In *Proceedings of the 33rd Annual International Conference on Salish and Neighboring Languages (ICSNL)*. University of British Columbia.
- Hess, T. and Bates, D. (2004). Lushootseed applicatives and their ilk. In Gertds, D. B. and Matthewson, L., editors, *Studies in Salish Linguistics in honor of M. Dale Kinkade*, number 17 in University of Montana Occasional Papers in Linguistics. University of Montana, Missoula, MT.
- Hess, T. and Hilbert, V. (1995a). *Lushootseed Book 1; The language of the Skagit, Nisqually, and other tribes of the Puget Sound. An Introduction*. Lushootseed Press, Seattle.
- Hess, T. and Hilbert, V. (1995b). *Lushootseed Book 2 (Advanced Lushootseed)*. Lushootseed Press, Seattle.
- Hess, T. and van Eijk, J. (1985). Noun and Verb in Salishan. In *20th International Conference on Salish and Neighboring Languages*, pages 97–104.
- Hess, T. M. (1967). *Snohomish Grammatical Structure*. PhD thesis, University of Washington, Seattle.
- Hess, T. M. (1982). Traces of ‘abnormal’ speech in Lushootseed. *International Conference on Salish and Neighboring Languages*, 7.
- Hess, T. M. (1996). *The Documentation of Lushootseed Language and Literature*, pages 5–8. In Bierwert (1996).
- Hilbert, V. (1974). On transcribing the Metcalf tapes. In *Proceedings of the International Conference on Salish and Neighboring Languages*, volume 9.
- Hilbert, V. (1976). Fieldwork report. In *Proceedings of the International Conference on Salish and Neighboring Languages*, volume 11.
- Hilbert, V. (1986). When Chief Seattle (siʔaʔ) spoke in 1855. In *Proceedings of the International Conference on Salish and Neighboring Languages*, volume 21.
- Hilbert, V. (2007). Letter from Vi Hilbert regarding the usage of Lushootseed alphabet. <http://depts.washington.edu/lutteach/?p=52>. Accessed: 2018-03-11T00:00+00:00.
- Hilbert, V., Miller, J., and Zahir, Z. (2001). *sdaʔdaʔ gʷəʔ dibəʔ ləʂucid ʔacaciʔtalbixʷ/Puget Sound Geography: Original Manuscript from T. T. Waterman*. Lushootseed Press.

- Hinrichs, E. and Nakazawa, T. (1990). Subcategorization and VP structure in German. In Hughes, S. and Salmons, J., editors, *Proc. 3rd Symposium on Germanic Linguistics*. Benjamins.
- Hodges, W. (2018). Tarski's truth definitions. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, fall 2018 edition.
- Hopcroft, J. E., Motwani, R., and Ullman, J. D. (2006). *Introduction to Automata Theory, Languages, and Computation*. Addison-Wesley, Boston, 2nd edition.
- Hopper, P. J. and Thompson, S. A. (1980). Transitivity in grammar and discourse. *Language*, 56(2):251–299.
- Howell, K., Zamaraeva, O., and Bender, E. M. (2018). Nominalized clauses in the Grammar Matrix. In Müller, S. and Richter, F., editors, *Proceedings of the 25th International Conference on Head-Driven Phrase Structure Grammar, University of Tokyo*, pages 68–88, Stanford, CA. CSLI Publications.
- Hugo, R. (2010). Indigenous language education in Washington: Facts, attitudes and vitality. Master's thesis, University of Washington, Seattle.
- Hugo, R. (2019). Indigenous language revitalization in the Pacific Northwest. In Denham, K., editor, *Northwest Voices*. University of Washington Press.
- Hukari, T. E. (1976). Transitivity in Halkomelem. *International Conference on Salish and Neighboring Languages*, 11:69–119.
- Hulden, M. (2009). Foma: A finite-state compiler and library. In *Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics: Demonstrations Session, EACL '09*, pages 29–32, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Hulden, M. and Alegria, I. (2010). Creating language resources and applications using finite-state morphological grammars. In *Language Resources and Evaluation Conference*.
- Hunt, H. and Kaylor, F. C. (1917). *Washington, West of the Cascades, Volume I*. The S. J. Clarke Publishing Company, Chicago, Seattle, Tacoma.
- International Phonetic Association (1999). Cambridge University Press, Cambridge.
- Jackendoff, R. (1972). *Semantics Interpretations in Generative Grammar*. MIT Press, Cambridge.

Jelinek, E. and Demers, R. A. (1983). The agent hierarchy and voice in some Coast Salish languages. *International Journal of American Linguistics*, 49(2):167–185.

Johnson, C. D. (1972). *Formal Aspects of Phonological Description*. Mouton, The Hague.

Joshi, A. K. (1985). Tree adjoining grammars: How much context-sensitivity is required to provide reasonable structural descriptions? In Dowty, D. R., Karttunen, L., and Zwicky, A. M., editors, *Natural Language Parsing*, pages 206–250. Cambridge University Press.

Jurafsky, D. and Martin, J. H. (2009). *Speech and Language Processing*. Pearson, Upper Saddle River.

Kaplan, R. M. and Kay, M. (1994). Regular models of phonological rule systems. *Computational Linguistics*, 20(3):331–378.

Kay, M. (1979). Functional grammar. *Annual Meeting of the Berkeley Linguistics Society*, 5(0):142–158.

Kinkade, D. M. (1970). Review of Southern Puget Sound Salish: Phonology and morphology and Southern Puget Sound Salish: Texts place names and dictionary. *American Anthropologist*, 72(5):1163–1165.

Kinkade, D. M. (1976). The copula and negatives in inland Olympic Salish. *International Journal of American Linguistics*, 42(1):17–23.

Kinkade, D. M. (1983). Salish evidence against the universality of ‘noun’ and ‘verb’. *Lingua*, 60:25–40.

Kinkade, D. M. (1985). More on nasal loss on the northwest coast. *International Journal of American Linguistics*, 51(4):478–480.

Kinkade, D. M. (1990). Prehistory of Salishan languages. In *International Conference on Salish and Neighboring Languages*, volume 25, Vancouver.

Kleene, S. C. (1956). Representation of events in nerve nets and finite automata. In Shannon, C. and McCarthy, J., editors, *Automata Studies*, pages 3–41. Princeton University Press, Princeton, NJ.

Kluger, R. (2011). *The Bitter Waters of Medicine Creek*. Vintage Books, New York.

Kratzer, A. (2017). Situations in natural language semantics. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, winter 2017 edition.

- Krauss, M. (1992). The world's languages in crisis. *Language*, 68(1):4–10.
- Krieger, H.-U. and Schaffer, U. (1994). TDL—a type description language for HPSG; part 1: Overview. Technical Report DFKI-RR-94-37, Deutsches Forschungszentrum für Künstliche Intelligenz GmbH.
- Kroeber, P. D. (1999). *The Salish Language Family: Reconstructing Syntax*. Studies in the Anthropology of North American Indians. University of Nebraska Press, Lincoln.
- Kruijff, G.-J. M. (2002). Formal and computational aspects of dependency grammar: History and development of DG. From ESSLI 2002. Accessed: 2018-12-08T18:51:06+00:00.
- Kuipers, A. H. (1968). The categories verb-noun and transitive-intransitive in English and Squamish. *Lingua*, 21:610–628.
- Kuipers, A. H. (2002). *Salish Etymological Dictionary*. Number 16 in Occasional Papers in Linguistics. University of Montana, Missoula.
- Langen, T. C. S. (1996). *Annotator's Introduction*, pages 52–62. In Bierwert (1996).
- Levin, B. (1993). *English Verb Classes and Alternations: a preliminary investigation*. University of Chicago Press, Chicago.
- Lien, E. and Kouylekov, M. (2014). Entailment recognition using minimal recursion semantics. In *Proceedings of the 8th International Workshop on Semantic Evaluation SemEval 2014*, pages 699–703.
- Lonsdale, D. and Matsushita, H. (2011). Annotating and exploring Lushootseed morphosyntax. In *Proceedings of the International Conference on Salish and Neighboring Languages: volume 46*, volume 30 of *University of British Columbia Working Papers in Linguistics*.
- MacKinlay, A., Martinez, D., and Baldwin, T. (2009). Biomedical event annotation with crfs and precision grammars. In *Proceedings of the BioNLP 2009 Workshop Companion Volume for Shared Task*, pages 77–85.
- Margolis, S. E. and Liebowitz, S. (2000). Path dependence: Look, I understand it's a little too late. <http://www.utdallas.edu/~liebowit/palgrave/palpd.html>. Accessed: 2019-03-09T22:56+00:00.
- Marino, C. (1990). History of Western Washington since 1846. In Suttles, W., editor, *Handbook of North American Indians Volume 7, Northwest Coast*, volume 7 of *Handbook of North American Indians*, pages 169–179. Smithsonian Institution.

Maxwell, M. (2012). Electronic grammars and reproducible research. In Nordhoff, S. and Poggeman, K.-L. G., editors, *Electronic Grammaticography*, number 4 in Language Documentation and Conservation Special Publication, pages 207–235. University of Hawaii Press, Honolulu.

McCarthy, J. (1960). Recursive functions of symbolic expressions and their computation by machine, part i. *Communications of the ACM*, 3(7):184–195.

McDonnell, J. A. (1991). *The Dispossession of the American Indian, 1887-1934*. Indiana University Press, Bloomington and Indianapolis.

Meeker, E. (1905). *Pioneer Reminiscences of Puget Sound: the tragedy of Leschi*. Lowman and Hanford Stationery and Printing Company, Seattle.

Meriam, L., Brown, R. A., Cloud, H. R., Dale, E. E., Duke, E., Edwards, H. R., McKenzie, F. A., Mark, M. L., Ryan, W. C. J., and Spillman, W. J. (1928). The problem of indian administration. Text available online: <https://www.narf.org/nill/resources/meriam.html>.

Miller, J. (2016). *George Gibbs Northwest Array: full reports, place names, word list, artifact names and guide*. Lushootseed Press.

Miller, J. (n.d.). U.S. archives with Lushootseed materials. Unpublished Manuscript by Lushootseed Research.

Miller, P. H. (1992). *Clitics and Constituents in Phrase Structure Grammar*. Garland, New York, London.

Mithun, M. (1997). Lexical affixes and morphological typology. In Bybee, J., Haiman, J., and Thompson, S. A., editors, *Essays on Language Function and Language Type*, pages 357–371. John Benjamins Publishing Company, Amsterdam.

Müller, S. (2013). The CoreGram project: A brief overview and motivation. In Duchier, D. and Parmentier, Y., editors, *Proceedings of the Workshop on High-level Methodologies for Grammar Engineering (HMGE 2013), Düsseldorf*, pages 93–104.

Müller, S. (2015). The CoreGram project: Theoretical linguistics, theory development and verification. *Journal of Language Modelling*, 3(1):21–86.

Nielsen, E. K. (2018). Modeling adnominal possession in the LinGO Grammar Matrix. Master's thesis, University of Washington, Seattle.

Oepen, S., Bender, E. M., Callmeier, U., Flickinger, D., and Siegel, M. (2002). Parallel distributed grammar engineering for practical applications. In *Proceedings of the Workshop on Grammar Engineering and Evaluation, COLING 2002*, Taipei, Taiwan.

Oepen, S. and Flickinger, D. P. (1998). Towards systematic grammar profiling. Test suite technology ten years after. *Journal of Computer Speech and Language*, 12 # 4 (Special Issue on Evaluation):411 – 436.

Oepen, S. and Lønning, J. T. (2006). Discriminant-based MRS banking. In Calzolari, N., Choukri, K., Gangemi, A., Maegaard, B., Mariani, J., Odijk, J., and Tapias, D., editors, *Proceedings of the Fifth International Conference on Language Resources and Evaluation, LREC 2006*, pages 1250–1255, Genoa, Italy.

Oepen, S., Netter, K., Bauer, J., Fettig, T., Klein, J., and Oberhauser, F. (1998). The tsnlp database - from tsct to tsdb. Technical Report LRE-62-089:D-WP6.1, Deutsches Forschungszentrum für Künstliche Intelligenz GmbH.

Packard, W. (2014). Uw-mrs: Leveraging a deep grammar for robotic spatial commands. In *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, pages 812–816.

Packard, W., Bender, E. M., Read, J., Oepen, S., and Drīdan, R. (2014). Simple negation scope resolution through deep parsing: A semantic solution to a semantic problem. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 69–78, Baltimore, Maryland. Association for Computational Linguistics.

Parsons, T. (1980). Modifiers and quantifiers in natural language. *Canadian Journal of Philosophy, Supplementary Volume*, 6:29.

Parsons, T. (1990). *Events in the semantics of English: a study in subatomic semantics*. MIT Press, Cambridge, MA.

Parsons, T. (1995). Thematic relations and arguments. *Linguistic Inquiry*, 26(4):625–662.

Payton, C. (2000). King county historical bibliography, part 02: Arthur c. ballard’s bibliography on the puget sound salish. <http://www.historylink.org/File/7143>. HistoryLink.org Essay 7143.

Pearson, J. D. (2003). Lewis Cass and the politics of disease: The Indian Vaccination Act of 1832. *Wicazo Sa Review*, 18(2):9–35.

Pollard, C. and Sag, I. A. (1987). *Information-based Syntax and Semantics, Volume 1: Fundamentals*. Cambridge University Press.

Pollard, C. and Sag, I. A. (1994). *Head-Driven Phrase Structure Grammar*. University of Chicago Press, London.

Poornima, S. and Koenig, J.-P. (2009). Hindi aspectual complex predicates. In Muller, S., editor, *Proceedings of the 16th International Conference on Head-Driven Phrase Structure Grammar*, pages 276–296. Stanford: CSLI Publications.

Pouson, L. (2011). Meta-modeling of tense and aspect in a cross-linguistic grammar engineering platform. *University of Washington Working Papers in Linguistics*, 28.

Prucha, F. P. (1984). *The Great Father: The United States Government and the American Indians*. University of Nebraska Press, Lincoln and London.

Puffert, D. J. (1999). Path dependence in Economic history. <https://web.archive.org/web/20060113120201/http://www.eh.net/encyclopedia/article/puffert.path.dependence>. Accessed: 2019-03-09T22:51+00:00.

Pāṇini (1897). *The Ashtadhyayi*. Sindhu Charan Bose, The Panini Office.

Ransom, J. E. (1945). Notes on Duwamish phonology and morphology. *International Journal of American Linguistics*, 11(4):204–210.

Rasmussen, P. (2014). What happened to the Steh-chass people? <https://oly-wa.us/EdibleForestGardens/PDF/HistoryOfSteh-chassPeople.pdf>. Accessed: 2019-09-15T00:11+00:00.

Reddic, S. M. and Collins, C. C. (2005). Medicine creek to Fox island: Cadastral scams and contested domains. *Oregon Historical Quarterly*, 106:374–397.

Riddle, M. (2009). St. Anne's Mission Church of Tulalip celebrates 100 years of service on December 15, 1957. <http://www.historylink.org/File/9049>. HistoryLink.org Essay 9049, accessed 2019-08-09T17:59+00:00.

Rosch, E. and Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, 8:382–439.

Sag, I. A. (2010). English filler-gap constructions. *Language*, 86:486–545.

Sag, I. A., Wasow, T., and Bender, E. M. (2003). *Syntactic Theory: A formal introduction*. CSLI, Stanford. Second Edition.

Saleem, S. and Bender, E. M. (2010). Argument optionality in the LinGO Grammar Matrix. In *Coling 2010: Posters*, pages 1068–1076, Beijing, China. Coling 2010 Organizing Committee.

Sheikh, A. and Collins, B. D. (2009). The waterlines project. Burke Museum of Natural History and Culture.

Shieber, S. (1985). Evidence against the context-freeness of natural language. *Linguistics and Philosophy*, 8(3):333–343.

Shieber, S. M. (1986). *An Introduction to Unification-Based Approaches to Grammar*. CSLI Publications, Stanford University, Stanford.

siastənu (2005). *ǰǰčusədə? ?ə siastənu: Gram Ruth Sehome Shelton; The Wisdom of a Tulalip Elder*. Lushootseed Press. A collection of materials recorded by Leon Metcalf. Includes an interview of Metcalf by Miller, with an introduction about Metcalf by Miller; Vi Hilbert and Jay Miller, eds.

Siegel, M. and Bender, E. M. (2002). Efficient deep processing of Japanese. In *Proceedings of the 3rd Workshop on Asian Language Resources and International Standardization at the 19th International Conference on Computational Linguistics*, Taipei, Taiwan.

Smith, B., Avis, D., Taylor, M., Perron, A., and Johnson, D. (2003). BeerXML: An XML standard for beer brewing data version 1.0. <http://www.beerxml.com/beerxml.htm>. Accessed: 2019-05-05T18:23+00:00.

Snyder, W. A. (1957). *A Phonemic and Morphological Analysis of Southern Puget Sound Salish*. PhD thesis, University of Washington, Seattle.

Snyder, W. A. (1968a). *Southern Puget Sound Salish: Phonology and Morphology*. Sacramento Anthropology Society, Sacramento State College, Sacramento.

Snyder, W. A. (1968b). *Southern Puget Sound Salish: Texts, Place names and Dictionary*. Sacramento Anthropology Society, Sacramento State College, Sacramento.

Song, S. (2017). *Modeling Information Structure in a Cross linguistic Perspective*. Language Science Press, Berlin.

Speaks, J. (2014). Theories of meaning. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, summer 2014 edition.

Stabler, E. (2004). Varieties of crossing dependencies: structure dependence and mild context sensitivity. *Cognitive Science*, 28(5):699–720.

Stassen, L. (2013). Zero copula for predicate nominals. In Dryer, M. S. and Haspelmath, M., editors, *The World Atlas of Language Structures Online*. Max Planck Institute for Evolutionary Anthropology, Leipzig.

Stevens, J. A. (1873). *A Memorial of George Gibbs*. New York.

Suttles, W. and Lane, B. (1990). Southern Coast Salish. In Suttles, W., editor, *Handbook of North American Indians Volume 7, Northwest Coast*, volume 7 of *Handbook of North American Indians*, pages 485–502. Smithsonian Institution.

Swadesh, M. (1950). Salish internal relationships. *International Journal of American Linguistics*, 16(4):157–167.

TEI Consortium (2019). Guidelines for electronic text encoding and interchange 3.5.0. 2019-01-29. Accessed 2019-05-05T18:10+00:00.

Tesnière, L. (1959). *Éléments de Syntaxe Structurale*. Klincksieck, Paris, France.

Thompson, L. C. (1979). Salishan and the Northwest. In Campbell, L. and M., M., editors, *The Languages of Native America: Historical and Comparative Assessment*, pages 692–765, Austin. University of Texas Press.

Thompson, L. C. and Kinkade, M. D. (1990). Languages. In Suttles, W., editor, *Handbook of North American Indians Volume 7, Northwest Coast*, volume 7 of *Handbook of North American Indians*, pages 30–51. Smithsonian Institution.

Thompson, L. C. and Thompson, T. M. (1980). Thompson Salish // -xi // . *International Journal of American Linguistics*, 46(1):27–32.

Tseng, J. L. (2003). EDGE features and French liaison. In Kim, J.-B. and Wechsler, S., editors, *The Proceedings of the 9th International Conference on Head-Driven Phrase Structure Grammar*, pages 313–333. Stanford: CSLI Publications.

Tweddell, C. E. (1950). The Snoqualmie-Duwamish dialects of Puget Sound Coast Salish. *University of Washington Publications in Anthropology*, 12(1):1–78.

United States Government Printing Office (1850). U.S. statutes at large, volume 9. Thirty-First Congress, Sess. I. Ch. 76.

Urbanczyk, S. (1996). *Patterns of Reduplication in Lushootseed*. PhD thesis, University of Massachusetts.

U.S. Board on Geographic Names Domestic Names Committee (2009). U.S. board on geographic names domestic names committee, seven hundred and fourteenth meeting.

van Eijk, J. (2009). Obituary: Thomas M. Hess. [url=https://linguistlist.org/issues/20/20-3455.html](https://linguistlist.org/issues/20/20-3455.html). Accessed: 2018-05-20T00:00+00:00.

van Eijk, J. and Hess, T. (1986). Noun and Verb in Salish. *Lingua*, 69(4):319–331.

van Eijk, J. and Williams, L., editors (1981). *Cuystwi Malh Ucwalmicwts: Lillooet Legends and Stories*. Ts'zil Publishers, Mount Currie, B.C.

Waterman, T. T. (1922). The geographical names used by Indians of the Pacific Coast. *Geographical Review*, 12(2):175–194.

Watson, K. G. (1999). Arthur Ballard, and the mythology of Southern Puget Sound. *White River Journal: a newsletter of the White River Valley Museum*.

Williams, S. (2013). Establishing a permanent peace: Civilizing and subjugating Native Americans during the peace policy era (1869-1877). Master's thesis, California State University, Stanislaus.

Wilma, D. (2000). Arthur Ballard records and translates native american legends of Puget Sound in 1916. <http://www.historylink.org/File/2563>. HistoryLink.org Essay 2563.

Wilma, D. (2003). Stevens, Isaac Ingalls (1818 – 1862). <http://www.historylink.org/File/5314>. HistoryLink.org Essay 5314.

Xia, F. and Lewis, W. (2009). Applying NLP technologies to the collection and enrichment of language data on the web to aid linguistic research. In *Proceedings of the EACL 2009 Workshop on Language Technology and Resources for Cultural Heritage, Social Sciences, Humanities, and Education (LaTeCH – SHELTeCH 2009)*, pages 51–59, Athens, Greece. Association for Computational Linguistics.

Yao, X., Bouma, G., and Zhang, Y. (2012). Semantics-based question generation and implementation. *Dialogue & Discourse*, 3(2):11–42.

Yoder, J. (2004). Hilbert, Vi. <http://www.historylink.org/File/7130>. Accessed: 2019-02-17T22:59+00:00.

Zahir, Z. (2016). Lushootseed lives. Unpublished, presented at Lushootseed Language conference 2016.

Zahir, Z. (2018). *Elements of Lushootseed Grammar in Discourse Perspective*. PhD thesis, University of Oregon, Eugene.

Zahir, Z. . (2009). A Lushootseed analysis of a 1877 dictionary by George Gibbs. <http://www.puyalluptriballanguage.org/resources/GibbsFinal4.pdf>. Accessed: 2019-03-13T00:00+00:00.

Łukasiewicz Jan (1970). On three-valued logic. In Borkowski, L., editor, *Selected works by Jan Łukasiewicz*, pages 87–88. North–Holland, Amsterdam.

## Appendix A

### SCRIPTS

This is the script used to generate the FST lexica from the XML source files. See Chapter 5 for context.

```
#!/usr/bin/env python3
fstgen
Author: Joshua Crwoyey
#
Generates .lexc files for Lushootseed Morphotactics based on an input
lexdb.xml file
Requires two command line arguments: the xml lexdb file and an output
directory

import sys
import os
import codecs

class FSTFile:

 def __init__(self, name, multicharsymbols, lexica):
 self.name = name
 self.multicharsymbols = multicharsymbols
 # lexica will be a list of dicts,
 # each dict is a lexicon with two keys: items, sub-dicts
 self.lexica = lexica

 def write_self(self, outdir):
 f = codecs.open(outdir + "/" + self.name, mode="w", encoding="utf-8")
 if not self.multicharsymbols.strip() == "":
 f.write("Multichar_Symbols {}\n\n".format(self.multicharsymbols))
 for l in self.lexica:
 f.write("LEXICON {}\n".format(l['name']))
 for i in l['items']:
 f.write("{}\n".format(i))
```

```

 if 'hasnull' in l and l['hasnull'] == True:
 f.write("#;\n")
 for sl in l['sublexica']:
 f.write("{};\n".format(sl.attrib['name']))
 f.write("\n")

def find_fstf_by_name(l ,n):
 for fstf in l:
 if fstf.name == n:
 return fstf
 print("can't find an fstf by name: "+n)
 sys.exit(1)

def usage():
 print("./fstgen LEXDBFILE FSTDIR")

check args
if len(sys.argv) != 3:
 usage()
 sys.exit(1)
elif not (os.path.isdir(sys.argv[2]) and os.path.isfile(sys.argv[1])):
 usage()
 sys.exit(1)
args okay, keep going...

from lxml import etree
lexdb = etree.parse(open(sys.argv[1])).getroot()

fstfs = []
create objects for each of the fst-files
for fstf in lexdb[0]:
 ms = ''
 if 'multichar-symbols' in fstf.attrib:
 ms = fstf.attrib['multichar-symbols']
 itr = fstf.iterchildren()
 lexica = []
 for i in itr:
 sublexica = []
 hasnull = False
 if i.findall('lexicon') != []:
 for sl in i.findall('lexicon'):
 sublexica.append(sl)

```

```

 if 'hasnull' in i.attrib:
 hasnull = True
 lexica.append({'name': i.attrib['name'],
 'items': [],
 'hasnull': hasnull,
 'sublexica': sublexica})
 fstfs.append(FSTFile(fstf.attrib['name'], ms, lexica))

for item in lexdb[1]:
 mp = item.find('mp')
 fstf = find_fstf_by_name(fstfs, mp.get('filename'))
 ln = mp[0].text
 for l in fstf.lexica:
 if l['name'] == ln:
 # prepare textual append
 uttext = mp.find('ut').text
 lttext = mp.find('lt').text
 if "radical" in mp.find('ut').attrib:
 # consider adding an end-radical symbol
 (i, j) = mp.find('ut').attrib['radical'].split(":")
 uttext = "\v{}}".format(uttext[0:int(i)], uttext[int(i):])
 if "radical" in mp.find('lt').attrib:
 # consider adding an end-radical symbol here too
 (i, j) = mp.find('lt').attrib['radical'].split(":")
 lttext = "\v{}}".format(lttext[0:int(i)], lttext[int(i):])

 if "flags" in mp.find('ut').attrib:
 uttext = ''.join(mp.find('ut').attrib['flags'].split()) + uttext
 if "flags" in mp.find('lt').attrib:
 lttext = ''.join(mp.find('lt').attrib['flags'].split()) + lttext

 text = ''
 if uttext == lttext:
 text = uttext
 else:
 text = "{}:{}".format(uttext, lttext)
 text += " #;"

 g = item.find('gloss').text
 if g is not None:
 text += " ! " + g

 l['items'].append(text)

```

```
for f in fstfs:
 f.write_self(sys.argv[2])
```

## Appendix B

**RESULTS FOR VALENCE INCREASING MORPHOLOGY TESTSUITE**

The following material is the formatted output of the syntactic grammar when run against the testsuite for valence increasing morphology. These materials are associated with the analysis and discussion of Chapter 7. Item numbers which are separated by 10s are traditional in grammar testsuites (because this allows the later addition of associated items without renumbering all subsequent material). Repeated item numbers in this appendix indicate that the grammar found more than one reading. Note that these sentences have also been published in full IGT form in Beck and Hess 2014.

In granting their permission to use these materials, the Tulalip Tribes Language Department asked me to print the following statement:

The lines from Thom Hess' transcription of Edward Sam's oral legacy are copyright by The Tulalip Tribes of Washington and are used in this work by permission of the Tulalip Tribes Lushootseed Department. The Department gratefully acknowledges Joshua Crowgey's valuable assistance in the improvement of the Department's website.

10 *ʔu-ʔuχʷ-yi-t ti čačas*

$\langle h_1, e_2 \{ \text{SF PROP-OR-QUES, E.ASPECT PERFECTIVE} \},$   
 $h_3: \text{"_ʔu}\chi\text{"_v\_go\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_4),$   
 $h_3: \text{"_yid\_x\_dat\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_5 \{ \text{PNG.PER 3RD} \}),$   
 $h_3: \text{"_t\_x\_ics\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_6),$   
 $h_7: \text{"_ti\_q\_unique\_rel"}(\text{ARG0: } x_5, \text{RSTR: } h_8, \text{BODY: } h_9),$   
 $h_{10}: \text{"_}\acute{c}\acute{a}\acute{c}\acute{a}\text{"_n\_child\_rel"}(\text{ARG0: } x_5)$   
 $\{ h_8 =_q h_{10} \} \rangle$

20 *ʔu-ʔuχ<sup>w</sup>-tx<sup>w</sup>-yi-t ti čáčas*

$\langle h_1, e_2\{\text{SF PROP-OR-QUES}\},$   
 $h_3: \text{"}_\text{ʔu}\chi^w\text{"}_v\text{"}_\text{go\_rel"}$ (ARG0:  $e_4\{\text{E.ASPECT PERFECTIVE}\}$ , ARG1:  $x_5$ ),  
 $h_3: \text{"}_\text{txw}_x\text{"}_\text{ecs\_rel"}$ (ARG0:  $e_4$ , ARG1:  $x_6$ ),  
 $h_7: \text{"}_\text{yid}_x\text{"}_\text{dat\_rel"}$ (ARG0:  $e_2$ , ARG1:  $x_8\{\text{PNG.PER 3RD}\}$ ),  
 $h_9: \text{"}_\text{ti}_q\text{"}_\text{unique\_rel"}$ (ARG0:  $x_8$ , RSTR:  $h_{10}$ , BODY:  $h_{11}$ ),  
 $h_{12}: \text{"}_\text{čáčas}_n\text{"}_\text{child\_rel"}$ (ARG0:  $x_8$ )  
 $\{ h_{10} =_q h_{12} \}$

30 *ʔu-k<sup>w</sup>ada-yi-t tsi čáčas*

$\langle h_1, e_2\{\text{SF PROP-OR-QUES, E.ASPECT PERFECTIVE}\},$   
 $h_3: \text{"}_\text{k}^w\text{ada}_v\text{"}_\text{take/grab\_rel"}$ (ARG0:  $e_2$ , ARG1:  $x_4$ ),  
 $h_3: \text{"}_\text{yid}_x\text{"}_\text{dat\_rel"}$ (ARG0:  $e_2$ , ARG1:  $x_5\{\text{PNG.PER 3RD, PNG.GEND FEM}\}$ ),  
 $h_3: \text{"}_\text{t}_x\text{"}_\text{ics\_rel"}$ (ARG0:  $e_2$ , ARG1:  $x_6$ ),  
 $h_7: \text{"}_\text{tsi}_q\text{"}_\text{unique.F\_rel"}$ (ARG0:  $x_5$ , RSTR:  $h_8$ , BODY:  $h_9$ ),  
 $h_{10}: \text{"}_\text{čáčas}_n\text{"}_\text{child\_rel"}$ (ARG0:  $x_5$ )  
 $\{ h_8 =_q h_{10} \}$

40 *ʔu-k<sup>w</sup>ada-yi-t ʔə ti ʔaʔχ*

$\langle h_1, e_2\{\text{SF PROP-OR-QUES, E.ASPECT PERFECTIVE}\},$   
 $h_3: \text{"}_\text{k}^w\text{ada}_v\text{"}_\text{take/grab\_rel"}$ (ARG0:  $e_2$ , ARG1:  $x_4\{\text{PNG.PER 3RD}\}$ ),  
 $h_3: \text{"}_\text{yid}_x\text{"}_\text{dat\_rel"}$ (ARG0:  $e_2$ , ARG1:  $i_5$ ),  
 $h_3: \text{"}_\text{t}_x\text{"}_\text{ics\_rel"}$ (ARG0:  $e_2$ , ARG1:  $x_6$ ),  
 $h_7: \text{"}_\text{ti}_q\text{"}_\text{unique\_rel"}$ (ARG0:  $x_4$ , RSTR:  $h_8$ , BODY:  $h_9$ ),  
 $h_{10}: \text{"}_\text{ʔa}\text{ʔ}\chi_n\text{"}_\text{platter\_rel"}$ (ARG0:  $x_4$ )  
 $\{ h_8 =_q h_{10} \}$

50 *ʔu-kʷəda-yi-t-b ʔə tsi luλ'*

$\langle h_1, e_2 \{E.ASPECT PERFECTIVE\},$   
 $h_3: \text{"_kʷəda\_v\_take/grab\_rel"}(ARG0: e_2, ARG1: x_4),$   
 $h_3: \text{"_yid\_x\_dat\_rel"}(ARG0: e_2, ARG1: i_5),$   
 $h_3: \text{"_t\_x\_ics\_rel"}(ARG0: e_2, ARG1: x_6 \{PNG.PER 3RD, PNG.GEND FEM\}),$   
 $h_7: \text{"_tsi\_q\_unique.F\_rel"}(ARG0: x_6, RSTR: h_8, BODY: h_9),$   
 $h_{10}: \text{"_luλ'\_n\_older-person\_rel"}(ARG0: x_6)$   
 $\{ h_8 =_q h_{10} \} \rangle$

50 *ʔu-kʷəda-yi-t-b ʔə tsi luλ'*

$\langle h_1, e_2 \{E.ASPECT PERFECTIVE\},$   
 $h_3: \text{"_kʷəda\_v\_take/grab\_rel"}(ARG0: e_2, ARG1: x_4 \{PNG.PER 3RD, PNG.GEND FEM\}),$   
 $h_3: \text{"_yid\_x\_dat\_rel"}(ARG0: e_2, ARG1: i_5),$   
 $h_3: \text{"_t\_x\_ics\_rel"}(ARG0: e_2, ARG1: i_6),$   
 $h_7: \text{"_tsi\_q\_unique.F\_rel"}(ARG0: x_4, RSTR: h_8, BODY: h_9),$   
 $h_{10}: \text{"_luλ'\_n\_older-person\_rel"}(ARG0: x_4)$   
 $\{ h_8 =_q h_{10} \} \rangle$

60 *ʔu-ʔab-yi-t ti sqʷəbayʔ*

$\langle h_1, e_2 \{SF PROP-OR-QUES, E.ASPECT PERFECTIVE\},$   
 $h_3: \text{"_ʔab\_extend\_rel"}(ARG0: e_2, ARG1: x_4),$   
 $h_3: \text{"_yid\_x\_dat\_rel"}(ARG0: e_2, ARG1: x_5 \{PNG.PER 3RD\}),$   
 $h_3: \text{"_t\_x\_ics\_rel"}(ARG0: e_2, ARG1: x_6),$   
 $h_7: \text{"_ti\_q\_unique\_rel"}(ARG0: x_5, RSTR: h_8, BODY: h_9),$   
 $h_{10}: \text{"_sqʷəbayʔ\_n\_dog\_rel"}(ARG0: x_5)$   
 $\{ h_8 =_q h_{10} \} \rangle$

70 *ʔu-ʔab-yi-t ʔə ti šaw*

$\langle h_1, e_2 \{ \text{SF PROP-OR-QUES, E.ASPECT PERFECTIVE} \},$   
 $h_3: \text{"_ʔab\_extend\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_4 \{ \text{PNG.PER 3RD} \}),$   
 $h_3: \text{"_yid\_x\_dat\_rel"}(\text{ARG0: } e_2, \text{ARG1: } i_5),$   
 $h_3: \text{"_t\_x\_ics\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_6),$   
 $h_7: \text{"_ti\_q\_unique\_rel"}(\text{ARG0: } x_4, \text{RSTR: } h_8, \text{BODY: } h_9),$   
 $h_{10}: \text{"_šaw\_n\_bone\_rel"}(\text{ARG0: } x_4)$   
 $\{ h_8 =_q h_{10} \} \rangle$

80 *ʔu-ʔab-yi-t-b ʔə ti čačas*

$\langle h_1, e_2 \{ \text{E.ASPECT PERFECTIVE} \},$   
 $h_3: \text{"_ʔab\_extend\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_4),$   
 $h_3: \text{"_yid\_x\_dat\_rel"}(\text{ARG0: } e_2, \text{ARG1: } i_5),$   
 $h_3: \text{"_t\_x\_ics\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_6 \{ \text{PNG.PER 3RD} \}),$   
 $h_7: \text{"_ti\_q\_unique\_rel"}(\text{ARG0: } x_6, \text{RSTR: } h_8, \text{BODY: } h_9),$   
 $h_{10}: \text{"_čačas\_n\_child\_rel"}(\text{ARG0: } x_6)$   
 $\{ h_8 =_q h_{10} \} \rangle$

80 *ʔu-ʔab-yi-t-b ʔə ti čačas*

$\langle h_1, e_2 \{ \text{E.ASPECT PERFECTIVE} \},$   
 $h_3: \text{"_ʔab\_extend\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_4 \{ \text{PNG.PER 3RD} \}),$   
 $h_3: \text{"_yid\_x\_dat\_rel"}(\text{ARG0: } e_2, \text{ARG1: } i_5),$   
 $h_3: \text{"_t\_x\_ics\_rel"}(\text{ARG0: } e_2, \text{ARG1: } i_6),$   
 $h_7: \text{"_ti\_q\_unique\_rel"}(\text{ARG0: } x_4, \text{RSTR: } h_8, \text{BODY: } h_9),$   
 $h_{10}: \text{"_čačas\_n\_child\_rel"}(\text{ARG0: } x_4)$   
 $\{ h_8 =_q h_{10} \} \rangle$

90 *ʔu-ʔuχ<sup>w</sup>-tx<sup>w</sup>-yi-t ʔə ti sq<sup>w</sup>əbayʔ*

⟨  $h_1, e_2$  {SF PROP-OR-QUES},

$h_3$ : "\_ʔuχ<sup>w</sup>\_v\_go\_rel"(ARG0:  $e_4$ {E.ASPECT PERFECTIVE}, ARG1:  $x_5$ {PNG.PER 3RD}),

$h_3$ : "\_txw\_x\_ecs\_rel"(ARG0:  $e_4$ , ARG1:  $x_6$ ),

$h_7$ : "\_yid\_x\_dat\_rel"(ARG0:  $e_2$ , ARG1:  $i_8$ ),

$h_9$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_{10}$ , BODY:  $h_{11}$ ),

$h_{12}$ : "\_sq<sup>w</sup>əbayʔ\_n\_dog\_rel"(ARG0:  $x_5$ )

{  $h_{10} =_q h_{12}$  } ⟩

100 *ʔu-ʔuχ<sup>w</sup>-tx<sup>w</sup>-yi-t-b ʔə ti luχ<sup>ʔ</sup>*

⟨  $h_1, e_2$ ,

$h_3$ : "\_ʔuχ<sup>w</sup>\_v\_go\_rel"(ARG0:  $e_4$ {E.ASPECT PERFECTIVE}, ARG1:  $x_5$ ),

$h_3$ : "\_txw\_x\_ecs\_rel"(ARG0:  $e_4$ , ARG1:  $x_6$ {PNG.PER 3RD}),

$h_7$ : "\_yid\_x\_dat\_rel"(ARG0:  $e_2$ , ARG1:  $i_8$ ),

$h_9$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_6$ , RSTR:  $h_{10}$ , BODY:  $h_{11}$ ),

$h_{12}$ : "\_luχ<sup>ʔ</sup>\_n\_older-person\_rel"(ARG0:  $x_6$ )

{  $h_{10} =_q h_{12}$  } ⟩

100 *ʔu-ʔuχ<sup>w</sup>-tx<sup>w</sup>-yi-t-b ʔə ti luχ<sup>ʔ</sup>*

⟨  $h_1, e_2$ ,

$h_3$ : "\_ʔuχ<sup>w</sup>\_v\_go\_rel"(ARG0:  $e_4$ {E.ASPECT PERFECTIVE}, ARG1:  $x_5$ {PNG.PER 3RD}),

$h_3$ : "\_txw\_x\_ecs\_rel"(ARG0:  $e_4$ , ARG1:  $i_6$ ),

$h_7$ : "\_yid\_x\_dat\_rel"(ARG0:  $e_2$ , ARG1:  $i_8$ ),

$h_9$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_{10}$ , BODY:  $h_{11}$ ),

$h_{12}$ : "\_luχ<sup>ʔ</sup>\_n\_older-person\_rel"(ARG0:  $x_5$ )

{  $h_{10} =_q h_{12}$  } ⟩

110 *ʔu-ʔəy̌-dx<sup>w</sup>-yi-t ti čačas*

⟨  $h_1, e_2$ {SF PROP-OR-QUES},

$h_3$ :"\_ʔəy̌\_v\_found\_rel"(ARG0:  $e_4$ {E.ASPECT PERFECTIVE}, ARG1:  $x_5$ ),

$h_3$ :"\_dxw\_x\_dc\_rel"(ARG0:  $e_4$ , ARG1:  $x_6$ ),

$h_7$ :"\_yid\_x\_dat\_rel"(ARG0:  $e_2$ , ARG1:  $x_8$ {PNG.PER 3RD}),

$h_9$ :"\_ti\_q\_unique\_rel"(ARG0:  $x_8$ , RSTR:  $h_{10}$ , BODY:  $h_{11}$ ),

$h_{12}$ :"\_čačas\_n\_child\_rel"(ARG0:  $x_8$ )

{  $h_{10} =_q h_{12}$  } ⟩

120 *ʔu-ʔəy̌-dx<sup>w</sup>-yi-t ʔə ti sqwəbay?*

⟨  $h_1, e_2$ {SF PROP-OR-QUES},

$h_3$ :"\_ʔəy̌\_v\_found\_rel"(ARG0:  $e_4$ {E.ASPECT PERFECTIVE}, ARG1:  $x_5$ {PNG.PER 3RD}),

$h_3$ :"\_dxw\_x\_dc\_rel"(ARG0:  $e_4$ , ARG1:  $x_6$ ),

$h_7$ :"\_yid\_x\_dat\_rel"(ARG0:  $e_2$ , ARG1:  $i_8$ ),

$h_9$ :"\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_{10}$ , BODY:  $h_{11}$ ),

$h_{12}$ :"\_sqwəbay?\_n\_dog\_rel"(ARG0:  $x_5$ )

{  $h_{10} =_q h_{12}$  } ⟩

130 *ʔu-ʔəy̌-dx<sup>w</sup>-yi-t-b ʔə tsi luχ'*

⟨  $h_1, e_2$ ,

$h_3$ :"\_ʔəy̌\_v\_found\_rel"(ARG0:  $e_4$ {E.ASPECT PERFECTIVE}, ARG1:  $x_5$ ),

$h_3$ :"\_dxw\_x\_dc\_rel"(ARG0:  $e_4$ , ARG1:  $x_6$ {PNG.PER 3RD, PNG.GEND FEM}),

$h_7$ :"\_yid\_x\_dat\_rel"(ARG0:  $e_2$ , ARG1:  $i_8$ ),

$h_9$ :"\_tsi\_q\_unique.F\_rel"(ARG0:  $x_6$ , RSTR:  $h_{10}$ , BODY:  $h_{11}$ ),

$h_{12}$ :"\_luχ'\_n\_older-person\_rel"(ARG0:  $x_6$ )

{  $h_{10} =_q h_{12}$  } ⟩

130 *ʔu-ʔəy̌-dx<sup>w</sup>-yi-t-b ʔə tsi luʔ*

⟨  $h_1, e_2,$   
 $h_3: \text{"_ʔəy̌_v_found\_rel"}(\text{ARG0: } e_4\{\text{E.ASPECT PERFECTIVE}\},$   
 $\text{ARG1: } x_5\{\text{PNG.PER 3RD, PNG.GEND FEM}\})$ ,  
 $h_3: \text{"_dxw\_x\_dc\_rel"}(\text{ARG0: } e_4, \text{ARG1: } i_6),$   
 $h_7: \text{"_yid\_x\_dat\_rel"}(\text{ARG0: } e_2, \text{ARG1: } i_8),$   
 $h_9: \text{"_tsi\_q\_unique.F\_rel"}(\text{ARG0: } x_5, \text{RSTR: } h_{10}, \text{BODY: } h_{11}),$   
 $h_{12}: \text{"_luʔ\_n\_older-person\_rel"}(\text{ARG0: } x_5)$   
 $\{ h_{10} =_q h_{12} \} \rangle$

140 *ʔu-ʔab-yi-t-b ʔə ti luʔ ʔə ti sq<sup>w</sup>əbay?*

⟨  $h_1, e_2\{\text{E.ASPECT PERFECTIVE}\},$   
 $h_3: \text{"_ʔab\_extend\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_4\{\text{PNG.PER 3RD}\}),$   
 $h_3: \text{"_yid\_x\_dat\_rel"}(\text{ARG0: } e_2, \text{ARG1: } i_5),$   
 $h_3: \text{"_t\_x\_ics\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_6\{\text{PNG.PER 3RD}\}),$   
 $h_7: \text{"_ti\_q\_unique\_rel"}(\text{ARG0: } x_6, \text{RSTR: } h_8, \text{BODY: } h_9),$   
 $h_{10}: \text{"_luʔ\_n\_older-person\_rel"}(\text{ARG0: } x_6),$   
 $h_{11}: \text{"_ti\_q\_unique\_rel"}(\text{ARG0: } x_4, \text{RSTR: } h_{12}, \text{BODY: } h_{13}),$   
 $h_{14}: \text{"_sq<sup>w</sup>əbay?_n\_dog\_rel"}(\text{ARG0: } x_4)$   
 $\{ h_8 =_q h_{10}, h_{12} =_q h_{14} \} \rangle$

150 *ʔu-ʔab-yi-t-b ʔə ti luʔ ti čăčas*

⟨  $h_1, e_2\{\text{E.ASPECT PERFECTIVE}\},$   
 $h_3: \text{"_ʔab\_extend\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_4),$   
 $h_3: \text{"_yid\_x\_dat\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_5\{\text{PNG.PER 3RD}\}),$   
 $h_3: \text{"_t\_x\_ics\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_6\{\text{PNG.PER 3RD}\}),$   
 $h_7: \text{"_ti\_q\_unique\_rel"}(\text{ARG0: } x_6, \text{RSTR: } h_8, \text{BODY: } h_9),$   
 $h_{10}: \text{"_luʔ\_n\_older-person\_rel"}(\text{ARG0: } x_6),$   
 $h_{11}: \text{"_ti\_q\_unique\_rel"}(\text{ARG0: } x_5, \text{RSTR: } h_{12}, \text{BODY: } h_{13}),$   
 $h_{14}: \text{"_čăčas\_n\_child\_rel"}(\text{ARG0: } x_5)$   
 $\{ h_8 =_q h_{10}, h_{12} =_q h_{14} \} \rangle$

160 *ʔu-ʔab-yi-t-b ti čačas ʔə ti sqwəbay?*

⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},  
 $h_3$ : "\_ʔab\_extend\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),  
 $h_3$ : "yid\_x\_dat\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD}),  
 $h_3$ : "\_t\_x\_ics\_rel"(ARG0:  $e_2$ , ARG1:  $i_6$ ),  
 $h_7$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_8$ , BODY:  $h_9$ ),  
 $h_{10}$ : "\_čačas\_n\_child\_rel"(ARG0:  $x_5$ ),  
 $h_{11}$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_{12}$ , BODY:  $h_{13}$ ),  
 $h_{14}$ : "\_sqwəbay?\_n\_dog\_rel"(ARG0:  $x_4$ )  
 {  $h_8 =_q h_{10}$ ,  $h_{12} =_q h_{14}$  } ⟩

170 *ʔu-ʔab-yi-t čəd ti čačas ʔə ti sqwəbay?*

⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},  
 $h_3$ : "\_ʔab\_extend\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),  
 $h_3$ : "yid\_x\_dat\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD}),  
 $h_3$ : "\_t\_x\_ics\_rel"(ARG0:  $e_2$ , ARG1:  $x_6$  {PNG.PER 1ST, PNG.NUM SG}),  
 $h_7$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_8$ , BODY:  $h_9$ ),  
 $h_{10}$ : "\_čačas\_n\_child\_rel"(ARG0:  $x_5$ ),  
 $h_{11}$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_{12}$ , BODY:  $h_{13}$ ),  
 $h_{14}$ : "\_sqwəbay?\_n\_dog\_rel"(ARG0:  $x_4$ )  
 {  $h_8 =_q h_{10}$ ,  $h_{12} =_q h_{14}$  } ⟩

180 *ʔu-ʔab-yi-t-b čəd ʔə ti čačas ʔə ti sqwəbay?*

⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},  
 $h_3$ : "\_ʔab\_extend\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),  
 $h_3$ : "yid\_x\_dat\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 1ST, PNG.NUM SG}),  
 $h_3$ : "\_t\_x\_ics\_rel"(ARG0:  $e_2$ , ARG1:  $x_6$  {PNG.PER 3RD}),  
 $h_7$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_6$ , RSTR:  $h_8$ , BODY:  $h_9$ ),  
 $h_{10}$ : "\_čačas\_n\_child\_rel"(ARG0:  $x_6$ ),  
 $h_{11}$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_{12}$ , BODY:  $h_{13}$ ),  
 $h_{14}$ : "\_sqwəbay?\_n\_dog\_rel"(ARG0:  $x_4$ )  
 {  $h_8 =_q h_{10}$ ,  $h_{12} =_q h_{14}$  } ⟩

190 *ʔu-kwəda-yi-t čəx<sup>w</sup> ʔu tsi čačas ʔə ti kʷataq*

$\langle h_1, e_2 \{SF\ QUES, E.ASPECT\ PERFECTIVE\},$

$h_3: \text{"_kʷəda\_v\_take/grab\_rel"}(ARG0: e_2, ARG1: x_4 \{PNG.PER\ 3RD\}),$

$h_3: \text{"yid\_x\_dat\_rel"}(ARG0: e_2, ARG1: x_5 \{PNG.PER\ 3RD, PNG.GEND\ FEM\}),$

$h_3: \text{"_t\_x\_ics\_rel"}(ARG0: e_2, ARG1: x_6 \{PNG.PER\ 2ND, PNG.NUM\ SG\}),$

$h_7: \text{"_tsi\_q\_unique.F\_rel"}(ARG0: x_5, RSTR: h_8, BODY: h_9),$

$h_{10}: \text{"_čačas\_n\_child\_rel"}(ARG0: x_5),$

$h_{11}: \text{"_ti\_q\_unique\_rel"}(ARG0: x_4, RSTR: h_{12}, BODY: h_{13}),$

$h_{14}: \text{"_kʷataq\_n\_mat\_rel"}(ARG0: x_4)$

$\{ h_8 =_q h_{10}, h_{12} =_q h_{14} \}$

200 *ʔu-kwəda-yi-t tsi čačas ʔə ti kʷataq*

$\langle h_1, e_2 \{SF\ PROP-OR-QUES, E.ASPECT\ PERFECTIVE\},$

$h_3: \text{"_kʷəda\_v\_take/grab\_rel"}(ARG0: e_2, ARG1: x_4 \{PNG.PER\ 3RD\}),$

$h_3: \text{"yid\_x\_dat\_rel"}(ARG0: e_2, ARG1: x_5 \{PNG.PER\ 3RD, PNG.GEND\ FEM\}),$

$h_3: \text{"_t\_x\_ics\_rel"}(ARG0: e_2, ARG1: x_6),$

$h_7: \text{"_tsi\_q\_unique.F\_rel"}(ARG0: x_5, RSTR: h_8, BODY: h_9),$

$h_{10}: \text{"_čačas\_n\_child\_rel"}(ARG0: x_5),$

$h_{11}: \text{"_ti\_q\_unique\_rel"}(ARG0: x_4, RSTR: h_{12}, BODY: h_{13}),$

$h_{14}: \text{"_kʷataq\_n\_mat\_rel"}(ARG0: x_4)$

$\{ h_8 =_q h_{10}, h_{12} =_q h_{14} \}$

210 *ʔu-kwəda-yi-t-b čəd ʔə tsi čačas ʔə ti kʷataq*

$\langle h_1, e_2 \{E.ASPECT\ PERFECTIVE\},$

$h_3: \text{"_kʷəda\_v\_take/grab\_rel"}(ARG0: e_2, ARG1: x_4 \{PNG.PER\ 3RD\}),$

$h_3: \text{"yid\_x\_dat\_rel"}(ARG0: e_2, ARG1: x_5 \{PNG.PER\ 1ST, PNG.NUM\ SG\}),$

$h_3: \text{"_t\_x\_ics\_rel"}(ARG0: e_2, ARG1: x_6 \{PNG.PER\ 3RD, PNG.GEND\ FEM\}),$

$h_7: \text{"_tsi\_q\_unique.F\_rel"}(ARG0: x_6, RSTR: h_8, BODY: h_9),$

$h_{10}: \text{"_čačas\_n\_child\_rel"}(ARG0: x_6),$

$h_{11}: \text{"_ti\_q\_unique\_rel"}(ARG0: x_4, RSTR: h_{12}, BODY: h_{13}),$

$h_{14}: \text{"_kʷataq\_n\_mat\_rel"}(ARG0: x_4)$

$\{ h_8 =_q h_{10}, h_{12} =_q h_{14} \}$

220 *ʔu-ləkʷ-yi-t čəd ti luλ'ʔə ti sʔuladxʷ*

⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},  
 $h_3$ : "\_ləkʷ\_v\_be.eaten\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),  
 $h_3$ : "yid\_x\_dat\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD}),  
 $h_3$ : "\_t\_x\_ics\_rel"(ARG0:  $e_2$ , ARG1:  $x_6$  {PNG.PER 1ST, PNG.NUM SG}),  
 $h_7$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_8$ , BODY:  $h_9$ ),  
 $h_{10}$ : "\_luλ'\_n\_older-person\_rel"(ARG0:  $x_5$ ),  
 $h_{11}$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_{12}$ , BODY:  $h_{13}$ ),  
 $h_{14}$ : "\_sʔuladxʷ\_n\_salmon\_rel"(ARG0:  $x_4$ )  
 {  $h_8 =_q h_{10}$ ,  $h_{12} =_q h_{14}$  } ⟩

230 *ʔu-ləkʷ-yi-t ti luλ'ʔə ti sʔuladxʷ*

⟨  $h_1, e_2$  {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 $h_3$ : "\_ləkʷ\_v\_be.eaten\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),  
 $h_3$ : "yid\_x\_dat\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD}),  
 $h_3$ : "\_t\_x\_ics\_rel"(ARG0:  $e_2$ , ARG1:  $x_6$ ),  
 $h_7$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_8$ , BODY:  $h_9$ ),  
 $h_{10}$ : "\_luλ'\_n\_older-person\_rel"(ARG0:  $x_5$ ),  
 $h_{11}$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_{12}$ , BODY:  $h_{13}$ ),  
 $h_{14}$ : "\_sʔuladxʷ\_n\_salmon\_rel"(ARG0:  $x_4$ )  
 {  $h_8 =_q h_{10}$ ,  $h_{12} =_q h_{14}$  } ⟩

240 *ʔu-ləkʷ-yi-t-b čəd ʔə ti luλ'ʔə ti sʔuladxʷ*

⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},  
 $h_3$ : "\_ləkʷ\_v\_be.eaten\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),  
 $h_3$ : "yid\_x\_dat\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 1ST, PNG.NUM SG}),  
 $h_3$ : "\_t\_x\_ics\_rel"(ARG0:  $e_2$ , ARG1:  $x_6$  {PNG.PER 3RD}),  
 $h_7$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_6$ , RSTR:  $h_8$ , BODY:  $h_9$ ),  
 $h_{10}$ : "\_luλ'\_n\_older-person\_rel"(ARG0:  $x_6$ ),  
 $h_{11}$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_{12}$ , BODY:  $h_{13}$ ),  
 $h_{14}$ : "\_sʔuladxʷ\_n\_salmon\_rel"(ARG0:  $x_4$ )  
 {  $h_8 =_q h_{10}$ ,  $h_{12} =_q h_{14}$  } ⟩

250 *ʔu-kʷəł-t ti qʷu?*

$\langle h_1, e_2 \{SF \text{ PROP-OR-QUES, E.ASPECT PERFECTIVE}\},$   
 $h_3: \text{"_kʷəł\_v\_pour\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_4 \{PNG.PER 3RD\}),$   
 $h_3: \text{"_t\_x\_ics\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_5),$   
 $h_6: \text{"_ti\_q\_unique\_rel"}(\text{ARG0: } x_4, \text{RSTR: } h_7, \text{BODY: } h_8),$   
 $h_9: \text{"_qʷu?_n\_water\_rel"}(\text{ARG0: } x_4)$   
 $\{ h_7 =_q h_9 \} \rangle$

260 *ʔu-kʷəł-dxʷ ti qʷu?*

$\langle h_1, e_2 \{SF \text{ PROP-OR-QUES, E.ASPECT PERFECTIVE}\},$   
 $h_3: \text{"_kʷəł\_v\_pour\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_4 \{PNG.PER 3RD\}),$   
 $h_3: \text{"_dxw\_x\_dc\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_5),$   
 $h_6: \text{"_ti\_q\_unique\_rel"}(\text{ARG0: } x_4, \text{RSTR: } h_7, \text{BODY: } h_8),$   
 $h_9: \text{"_qʷu?_n\_water\_rel"}(\text{ARG0: } x_4)$   
 $\{ h_7 =_q h_9 \} \rangle$

270 *ʔu-bəč-t ti čačas*

$\langle h_1, e_2 \{SF \text{ PROP-OR-QUES, E.ASPECT PERFECTIVE}\},$   
 $h_3: \text{"_bəč\_v\_fallen\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_4 \{PNG.PER 3RD\}),$   
 $h_3: \text{"_t\_x\_ics\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_5),$   
 $h_6: \text{"_ti\_q\_unique\_rel"}(\text{ARG0: } x_4, \text{RSTR: } h_7, \text{BODY: } h_8),$   
 $h_9: \text{"_čačas\_n\_child\_rel"}(\text{ARG0: } x_4)$   
 $\{ h_7 =_q h_9 \} \rangle$

280 *ʔu-bəč-dxʷ ti čačas*

$\langle h_1, e_2 \{SF \text{ PROP-OR-QUES, E.ASPECT PERFECTIVE}\},$   
 $h_3: \text{"_bəč\_v\_fallen\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_4 \{PNG.PER 3RD\}),$   
 $h_3: \text{"_dxw\_x\_dc\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_5),$   
 $h_6: \text{"_ti\_q\_unique\_rel"}(\text{ARG0: } x_4, \text{RSTR: } h_7, \text{BODY: } h_8),$   
 $h_9: \text{"_čačas\_n\_child\_rel"}(\text{ARG0: } x_4)$   
 $\{ h_7 =_q h_9 \} \rangle$

290 *ʔu-kwəda-t ti sqwəbay?*

⟨  $h_1, e_2$ {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 $h_3$ :"\_kwəda\_v\_take/grab\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ {PNG.PER 3RD}),  
 $h_3$ :"\_t\_x\_ics\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$ ),  
 $h_6$ :"\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ :"\_sqwəbay?\_n\_dog\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9$  } ⟩

300 *ʔu-kwəda-dxw ti sqwəbay?*

⟨  $h_1, e_2$ {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 $h_3$ :"\_kwəda\_v\_take/grab\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ {PNG.PER 3RD}),  
 $h_3$ :"\_dxw\_x\_dc\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$ ),  
 $h_6$ :"\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ :"\_sqwəbay?\_n\_dog\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9$  } ⟩

310 *ʔu-čəl-t ti luλ'*

⟨  $h_1, e_2$ {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 $h_3$ :"\_čəl\_v\_defeated\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ {PNG.PER 3RD}),  
 $h_3$ :"\_t\_x\_ics\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$ ),  
 $h_6$ :"\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ :"\_luλ'\_n\_older-person\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9$  } ⟩

320 *ʔu-čəl-dxw ti luλ'*

⟨  $h_1, e_2$ {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 $h_3$ :"\_čəl\_v\_defeated\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ {PNG.PER 3RD}),  
 $h_3$ :"\_dxw\_x\_dc\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$ ),  
 $h_6$ :"\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ :"\_luλ'\_n\_older-person\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9$  } ⟩

330 *ʔu-ʔəy̌-dx<sup>w</sup> čəd ti čáčas*

$\langle h_1, e_2 \{E.ASPECT PERFECTIVE\},$   
 $h_3: \text{"_ʔəy̌_v_found_rel"}(ARG0: e_2, ARG1: x_4 \{PNG.PER 3RD\}),$   
 $h_3: \text{"_dxw_x_dc_rel"}(ARG0: e_2, ARG1: x_5 \{PNG.PER 1ST, PNG.NUM SG\}),$   
 $h_6: \text{"_ti_q_unique_rel"}(ARG0: x_4, RSTR: h_7, BODY: h_8),$   
 $h_9: \text{"_čáčas_n_child_rel"}(ARG0: x_4)$   
 $\{ h_7 =_q h_9 \} \rangle$

340 *ʔu-ʔəy̌-dx<sup>w</sup> ti sq<sup>w</sup>əbay?*

$\langle h_1, e_2 \{SF PROP-OR-QUES, E.ASPECT PERFECTIVE\},$   
 $h_3: \text{"_ʔəy̌_v_found_rel"}(ARG0: e_2, ARG1: x_4 \{PNG.PER 3RD\}),$   
 $h_3: \text{"_dxw_x_dc_rel"}(ARG0: e_2, ARG1: x_5),$   
 $h_6: \text{"_ti_q_unique_rel"}(ARG0: x_4, RSTR: h_7, BODY: h_8),$   
 $h_9: \text{"_sq<sup>w</sup>əbay?_n_dog_rel"}(ARG0: x_4)$   
 $\{ h_7 =_q h_9 \} \rangle$

350 *ʔu-ʔəy̌-dx<sup>w</sup> čəx<sup>w</sup> ʔu ti sq<sup>w</sup>əbay?*

$\langle h_1, e_2 \{SF QUES, E.ASPECT PERFECTIVE\},$   
 $h_3: \text{"_ʔəy̌_v_found_rel"}(ARG0: e_2, ARG1: x_4 \{PNG.PER 3RD\}),$   
 $h_3: \text{"_dxw_x_dc_rel"}(ARG0: e_2, ARG1: x_5 \{PNG.PER 2ND, PNG.NUM SG\}),$   
 $h_6: \text{"_ti_q_unique_rel"}(ARG0: x_4, RSTR: h_7, BODY: h_8),$   
 $h_9: \text{"_sq<sup>w</sup>əbay?_n_dog_rel"}(ARG0: x_4)$   
 $\{ h_7 =_q h_9 \} \rangle$

360 *ʔu-ʔux<sup>w</sup>-c čəd ti sqig<sup>w</sup>ac*

$\langle h_1, e_2 \{E.ASPECT PERFECTIVE\},$   
 $h_3: \text{"_ʔux<sup>w</sup>_v_go_rel"}(ARG0: e_2, ARG1: x_4 \{PNG.PER 1ST, PNG.NUM SG\}),$   
 $h_3: \text{"_c_x_altv_rel"}(ARG0: e_2, ARG1: x_5 \{PNG.PER 3RD\}),$   
 $h_6: \text{"_ti_q_unique_rel"}(ARG0: x_5, RSTR: h_7, BODY: h_8),$   
 $h_9: \text{"_sqig<sup>w</sup>ac_n_deer_rel"}(ARG0: x_5)$   
 $\{ h_7 =_q h_9 \} \rangle$

370 *ʔu-čubə-c ʔu ti čačas*

⟨  $h_1, e_2$  {SF QUES, E.ASPECT PERFECTIVE},  
 $h_3$ : "\_čubə\_v\_go.inland\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ ),  
 $h_3$ : "\_c\_x\_altv\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD}),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_čačas\_n\_child\_rel"(ARG0:  $x_5$ )  
 $\{ h_7 =_q h_9 \}$  ⟩

380 *ʔu-bəč-dx<sup>w</sup> čəx<sup>w</sup> ʔu ti luλ'*

⟨  $h_1, e_2$  {SF QUES, E.ASPECT PERFECTIVE},  
 $h_3$ : "\_bəč\_v\_fallen\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),  
 $h_3$ : "\_dxw\_x\_dc\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 2ND, PNG.NUM SG}),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_luλ'\_n\_older-person\_rel"(ARG0:  $x_4$ )  
 $\{ h_7 =_q h_9 \}$  ⟩

390 *ʔu-k<sup>w</sup>əł-t čəd ti q<sup>w</sup>u?*

⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},  
 $h_3$ : "\_k<sup>w</sup>əł\_v\_pour\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),  
 $h_3$ : "\_t\_x\_ics\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 1ST, PNG.NUM SG}),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_q<sup>w</sup>u?\_n\_water\_rel"(ARG0:  $x_4$ )  
 $\{ h_7 =_q h_9 \}$  ⟩

400 *ʔu-čəl-dx<sup>w</sup> čəx<sup>w</sup> ʔu*

⟨  $h_1, e_2$  {SF QUES, E.ASPECT PERFECTIVE},  
 $h_3$ : "\_čəl\_v\_defeated\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ ),  
 $h_3$ : "\_dxw\_x\_dc\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 2ND, PNG.NUM SG})  
 $\{ \}$  ⟩

410 *ʔu-kʷəda-t ti čačas*

⟨  $h_1, e_2$  {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 $h_3$ : "\_kʷəda\_v\_take/grab\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),  
 $h_3$ : "\_t\_x\_ics\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$ ),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_čačas\_n\_child\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9$  } ⟩

420 *ʔu-kʷəda-dxʷ čaxʷ ʔu ti sʔuladxʷ*

⟨  $h_1, e_2$  {SF QUES, E.ASPECT PERFECTIVE},  
 $h_3$ : "\_kʷəda\_v\_take/grab\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),  
 $h_3$ : "\_dxw\_x\_dc\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 2ND, PNG.NUM SG}),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_sʔuladxʷ\_n\_salmon\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9$  } ⟩

430 *ʔu-bəč-t čəł tsi luλ'*

⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},  
 $h_3$ : "\_bəč\_v\_fallen\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD, PNG.GEND FEM}),  
 $h_3$ : "\_t\_x\_ics\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 1ST, PNG.NUM PL}),  
 $h_6$ : "\_tsi\_q\_unique.F\_rel"(ARG0:  $x_4$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_luλ'\_n\_older-person\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9$  } ⟩

440 *ʔu-ʔəy'-dxʷ tsi luλ'*

⟨  $h_1, e_2$  {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 $h_3$ : "\_ʔəy'\_v\_found\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD, PNG.GEND FEM}),  
 $h_3$ : "\_dxw\_x\_dc\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$ ),  
 $h_6$ : "\_tsi\_q\_unique.F\_rel"(ARG0:  $x_4$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_luλ'\_n\_older-person\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9$  } ⟩

450 *?u-?usil ti čičičič*

$\langle h_1, e_2 \{SF \text{ PROP-OR-QUES}, E.ASPECT \text{ PERFECTIVE}\},$   
 $h_3: \text{"_?usil\_v\_dive\_rel"}(ARG0: e_2, ARG1: x_4 \{PNG.PER \text{ 3RD}\}),$   
 $h_5: \text{"_ti\_q\_unique\_rel"}(ARG0: x_4, RSTR: h_6, BODY: h_7),$   
 $h_8: \text{"_čičičičw\_n\_osprey\_rel"}(ARG0: x_4)$   
 $\{ h_6 =_q h_8 \} \rangle$

460 *?u-?usil-s ti s?uladx<sup>w</sup>*

$\langle h_1, e_2 \{SF \text{ PROP-OR-QUES}, E.ASPECT \text{ PERFECTIVE}\},$   
 $h_3: \text{"_?usil\_v\_dive\_rel"}(ARG0: e_2, ARG1: x_4),$   
 $h_3: \text{"_c\_x\_altv\_rel"}(ARG0: e_2, ARG1: x_5 \{PNG.PER \text{ 3RD}\}),$   
 $h_6: \text{"_ti\_q\_unique\_rel"}(ARG0: x_5, RSTR: h_7, BODY: h_8),$   
 $h_9: \text{"_s?uladxw\_n\_salmon\_rel"}(ARG0: x_5)$   
 $\{ h_7 =_q h_9 \} \rangle$

470 *?u-təlawil ti sq<sup>w</sup>əbay?*

$\langle h_1, e_2 \{SF \text{ PROP-OR-QUES}, E.ASPECT \text{ PERFECTIVE}\},$   
 $h_3: \text{"_təlawil\_v\_run\_rel"}(ARG0: e_2, ARG1: x_4 \{PNG.PER \text{ 3RD}\}),$   
 $h_5: \text{"_ti\_q\_unique\_rel"}(ARG0: x_4, RSTR: h_6, BODY: h_7),$   
 $h_8: \text{"_sq<sup>w</sup>əbay?_n\_dog\_rel"}(ARG0: x_4)$   
 $\{ h_6 =_q h_8 \} \rangle$

480 *?u-təlawil-s ti sqig<sup>w</sup>ac*

$\langle h_1, e_2 \{SF \text{ PROP-OR-QUES}, E.ASPECT \text{ PERFECTIVE}\},$   
 $h_3: \text{"_təlawil\_v\_run\_rel"}(ARG0: e_2, ARG1: x_4),$   
 $h_3: \text{"_c\_x\_altv\_rel"}(ARG0: e_2, ARG1: x_5 \{PNG.PER \text{ 3RD}\}),$   
 $h_6: \text{"_ti\_q\_unique\_rel"}(ARG0: x_5, RSTR: h_7, BODY: h_8),$   
 $h_9: \text{"_sqig<sup>w</sup>ac\_n\_deer\_rel"}(ARG0: x_5)$   
 $\{ h_7 =_q h_9 \} \rangle$

490 *?u-təlil ti luλ'*

$\langle h_1, e_2 \{SF \text{ PROP-OR-QUES}, E.ASPECT \text{ PERFECTIVE}\},$   
 $h_3: \text{"_təlil\_v\_go.ashore\_rel"}(ARG0: e_2, ARG1: x_4 \{PNG.PER \text{ 3RD}\}),$   
 $h_5: \text{"_ti\_q\_unique\_rel"}(ARG0: x_4, RSTR: h_6, BODY: h_7),$   
 $h_8: \text{"_luλ'\_n\_older-person\_rel"}(ARG0: x_4)$   
 $\{ h_6 =_q h_8 \} \rangle$

500 *ʔu-ʔalil-s ti supqs*

$\langle h_1, e_2 \{SF \text{ PROP-OR-QUES}, E.ASPECT \text{ PERFECTIVE}\},$   
 $h_3: \text{"_ʔalil\_v\_go.ashore\_rel"}(ARG0: e_2, ARG1: x_4),$   
 $h_3: \text{"_c\_x\_altv\_rel"}(ARG0: e_2, ARG1: x_5 \{PNG.PER \text{ 3RD}\}),$   
 $h_6: \text{"_ti\_q\_unique\_rel"}(ARG0: x_5, RSTR: h_7, BODY: h_8),$   
 $h_9: \text{"_supqs\_n\_hair.seal[snohomish]\_rel"}(ARG0: x_5)$   
 $\{ h_7 =_q h_9 \} \rangle$

510 *ʔu-ʔčil tsi čačas*

$\langle h_1, e_2 \{SF \text{ PROP-OR-QUES}, E.ASPECT \text{ PERFECTIVE}\},$   
 $h_3: \text{"_ʔčil\_v\_arrive\_rel"}(ARG0: e_2, ARG1: x_4 \{PNG.PER \text{ 3RD}, PNG.GEND \text{ FEM}\}),$   
 $h_5: \text{"_tsi\_q\_unique.F\_rel"}(ARG0: x_4, RSTR: h_6, BODY: h_7),$   
 $h_8: \text{"_čačas\_n\_child\_rel"}(ARG0: x_4)$   
 $\{ h_6 =_q h_8 \} \rangle$

520 *ʔu-ʔčil-s ti luλ'*

$\langle h_1, e_2 \{SF \text{ PROP-OR-QUES}, E.ASPECT \text{ PERFECTIVE}\},$   
 $h_3: \text{"_ʔčil\_v\_arrive\_rel"}(ARG0: e_2, ARG1: x_4),$   
 $h_3: \text{"_c\_x\_altv\_rel"}(ARG0: e_2, ARG1: x_5 \{PNG.PER \text{ 3RD}\}),$   
 $h_6: \text{"_ti\_q\_unique\_rel"}(ARG0: x_5, RSTR: h_7, BODY: h_8),$   
 $h_9: \text{"_luλ'\_n\_older-person\_rel"}(ARG0: x_5)$   
 $\{ h_7 =_q h_9 \} \rangle$

530 *ʔu-x<sup>wi</sup>tił ti čačas*

$\langle h_1, e_2 \{SF \text{ PROP-OR-QUES}, E.ASPECT \text{ PERFECTIVE}\},$   
 $h_3: \text{"_x<sup>wi</sup>tił\_v\_fall.off\_rel"}(ARG0: e_2, ARG1: x_4 \{PNG.PER \text{ 3RD}\}),$   
 $h_5: \text{"_ti\_q\_unique\_rel"}(ARG0: x_4, RSTR: h_6, BODY: h_7),$   
 $h_8: \text{"_čačas\_n\_child\_rel"}(ARG0: x_4)$   
 $\{ h_6 =_q h_8 \} \rangle$

540 *ʔu-xʷit̪il-t ti čačas*

$\langle h_1, e_2 \{SF \text{ PROP-OR-QUES}, E.ASPECT \text{ PERFECTIVE}\},$   
 $h_3: \text{"_xʷit̪il\_v\_fall.off\_rel"}(ARG0: e_2, ARG1: x_4 \{PNG.PER \text{ 3RD}\}),$   
 $h_3: \text{"_t\_x\_ics\_rel"}(ARG0: e_2, ARG1: x_5),$   
 $h_6: \text{"_ti\_q\_unique\_rel"}(ARG0: x_4, RSTR: h_7, BODY: h_8),$   
 $h_9: \text{"_čačas\_n\_child\_rel"}(ARG0: x_4)$   
 $\{ h_7 =_q h_9 \} \rangle$

550 *ʔu-ct̪il ti sqigʷac*

$\langle h_1, e_2 \{SF \text{ PROP-OR-QUES}, E.ASPECT \text{ PERFECTIVE}\},$   
 $h_3: \text{"_cʔil\_v\_bleed\_rel"}(ARG0: e_2, ARG1: x_4 \{PNG.PER \text{ 3RD}\}),$   
 $h_5: \text{"_ti\_q\_unique\_rel"}(ARG0: x_4, RSTR: h_6, BODY: h_7),$   
 $h_8: \text{"_sqigʷac\_n\_deer\_rel"}(ARG0: x_4)$   
 $\{ h_6 =_q h_8 \} \rangle$

560 *ʔu-ct̪il-t ti sqigʷac*

$\langle h_1, e_2 \{SF \text{ PROP-OR-QUES}, E.ASPECT \text{ PERFECTIVE}\},$   
 $h_3: \text{"_cʔil\_v\_bleed\_rel"}(ARG0: e_2, ARG1: x_4 \{PNG.PER \text{ 3RD}\}),$   
 $h_3: \text{"_t\_x\_ics\_rel"}(ARG0: e_2, ARG1: x_5),$   
 $h_6: \text{"_ti\_q\_unique\_rel"}(ARG0: x_4, RSTR: h_7, BODY: h_8),$   
 $h_9: \text{"_sqigʷac\_n\_deer\_rel"}(ARG0: x_4)$   
 $\{ h_7 =_q h_9 \} \rangle$

570 *ʔu-λuxʷil ti sʔəʔəd*

$\langle h_1, e_2 \{SF \text{ PROP-OR-QUES}, E.ASPECT \text{ PERFECTIVE}\},$   
 $h_3: \text{"_λuxʷil\_v\_become.cold\_rel"}(ARG0: e_2, ARG1: x_4 \{PNG.PER \text{ 3RD}\}),$   
 $h_5: \text{"_ti\_q\_unique\_rel"}(ARG0: x_4, RSTR: h_6, BODY: h_7),$   
 $h_8: \text{"_sʔəʔəd\_n\_water\_rel"}(ARG0: x_4)$   
 $\{ h_6 =_q h_8 \} \rangle$

580 *ʔu-λux<sup>w</sup>il-t ti sʔəʔəd*

$\langle h_1, e_2 \{SF \text{ PROP-OR-QUES}, E.ASPECT \text{ PERFECTIVE}\},$   
 $h_3: \text{"_λux<sup>w</sup>il\_v\_become.cold\_rel"}(ARG0: e_2, ARG1: x_4 \{PNG.PER \text{ 3RD}\}),$   
 $h_3: \text{"_t\_x\_ics\_rel"}(ARG0: e_2, ARG1: x_5),$   
 $h_6: \text{"_ti\_q\_unique\_rel"}(ARG0: x_4, RSTR: h_7, BODY: h_8),$   
 $h_9: \text{"_sʔəʔəd\_n\_water\_rel"}(ARG0: x_4)$   
 $\{ h_7 =_q h_9 \} \rangle$

590 *ʔu-čcil ti čaləs*

$\langle h_1, e_2 \{SF \text{ PROP-OR-QUES}, E.ASPECT \text{ PERFECTIVE}\},$   
 $h_3: \text{"_čcil\_v\_become.red\_rel"}(ARG0: e_2, ARG1: x_4 \{PNG.PER \text{ 3RD}\}),$   
 $h_5: \text{"_ti\_q\_unique\_rel"}(ARG0: x_4, RSTR: h_6, BODY: h_7),$   
 $h_8: \text{"_čaləs\_n\_hand\_rel"}(ARG0: x_4)$   
 $\{ h_6 =_q h_8 \} \rangle$

600 *ʔu-čcil-t ti čaləs*

$\langle h_1, e_2 \{SF \text{ PROP-OR-QUES}, E.ASPECT \text{ PERFECTIVE}\},$   
 $h_3: \text{"_čcil\_v\_become.red\_rel"}(ARG0: e_2, ARG1: x_4 \{PNG.PER \text{ 3RD}\}),$   
 $h_3: \text{"_t\_x\_ics\_rel"}(ARG0: e_2, ARG1: x_5),$   
 $h_6: \text{"_ti\_q\_unique\_rel"}(ARG0: x_4, RSTR: h_7, BODY: h_8),$   
 $h_9: \text{"_čaləs\_n\_hand\_rel"}(ARG0: x_4)$   
 $\{ h_7 =_q h_9 \} \rangle$

610 *ʔu-ʔuχ<sup>w</sup> tsi luχ<sup>ʔ</sup>*

$\langle h_1, e_2 \{SF \text{ PROP-OR-QUES}, E.ASPECT \text{ PERFECTIVE}\},$   
 $h_3: \text{"_ʔuχ<sup>w</sup>\_v\_go\_rel"}(ARG0: e_2, ARG1: x_4 \{PNG.PER \text{ 3RD}, PNG.GEND \text{ FEM}\}),$   
 $h_5: \text{"_tsi\_q\_unique.F\_rel"}(ARG0: x_4, RSTR: h_6, BODY: h_7),$   
 $h_8: \text{"_luχ<sup>ʔ</sup>\_n\_older-person\_rel"}(ARG0: x_4)$   
 $\{ h_6 =_q h_8 \} \rangle$

620 *ʔu-ʔuχ<sup>w</sup>-tx<sup>w</sup> tsi luλ'*

$\langle h_1, e_2 \{SF \text{ PROP-OR-QUES}, E.ASPECT \text{ PERFECTIVE}\},$   
 $h_3: \text{"_ʔu}\chi^w\_v\_go\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_4 \{PNG.PER \text{ 3RD}, PNG.GEND \text{ FEM}\}),$   
 $h_3: \text{"_txw\_x\_ecs\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_5),$   
 $h_6: \text{"_tsi\_q\_unique.F\_rel"}(\text{ARG0: } x_4, \text{RSTR: } h_7, \text{BODY: } h_8),$   
 $h_9: \text{"_lu}\lambda'_n\_older\_person\_rel"}(\text{ARG0: } x_4)$   
 $\{ h_7 =_q h_9 \} \rangle$

630 *ʔu-ʔuχ<sup>w</sup>-c tsi luλ'*

$\langle h_1, e_2 \{SF \text{ PROP-OR-QUES}, E.ASPECT \text{ PERFECTIVE}\},$   
 $h_3: \text{"_ʔu}\chi^w\_v\_go\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_4),$   
 $h_3: \text{"_c\_x\_altv\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_5 \{PNG.PER \text{ 3RD}, PNG.GEND \text{ FEM}\}),$   
 $h_6: \text{"_tsi\_q\_unique.F\_rel"}(\text{ARG0: } x_5, \text{RSTR: } h_7, \text{BODY: } h_8),$   
 $h_9: \text{"_lu}\lambda'_n\_older\_person\_rel"}(\text{ARG0: } x_5)$   
 $\{ h_7 =_q h_9 \} \rangle$

640 *ʔu-ʔəλ'ti sq<sup>w</sup>əbay?*

$\langle h_1, e_2 \{SF \text{ PROP-OR-QUES}, E.ASPECT \text{ PERFECTIVE}\},$   
 $h_3: \text{"_ʔə}\lambda'_v\_come\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_4 \{PNG.PER \text{ 3RD}\}),$   
 $h_5: \text{"_ti\_q\_unique\_rel"}(\text{ARG0: } x_4, \text{RSTR: } h_6, \text{BODY: } h_7),$   
 $h_8: \text{"_sq}^w\text{əbay?}_n\_dog\_rel"}(\text{ARG0: } x_4)$   
 $\{ h_6 =_q h_8 \} \rangle$

650 *ʔu-ʔəλ'-tx<sup>w</sup> ti sq<sup>w</sup>əbay?*

$\langle h_1, e_2 \{SF \text{ PROP-OR-QUES}, E.ASPECT \text{ PERFECTIVE}\},$   
 $h_3: \text{"_ʔə}\lambda'_v\_come\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_4 \{PNG.PER \text{ 3RD}\}),$   
 $h_3: \text{"_txw\_x\_ecs\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_5),$   
 $h_6: \text{"_ti\_q\_unique\_rel"}(\text{ARG0: } x_4, \text{RSTR: } h_7, \text{BODY: } h_8),$   
 $h_9: \text{"_sq}^w\text{əbay?}_n\_dog\_rel"}(\text{ARG0: } x_4)$   
 $\{ h_7 =_q h_9 \} \rangle$

660 *ʔu-ʔəʎ'-c ti sqʷəbayʔ*

⟨  $h_1, e_2$  {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 $h_3$ : "\_ʔəʎ'\_v\_come\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ ),  
 $h_3$ : "\_c\_x\_altv\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD}),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_sqʷəbayʔ\_n\_dog\_rel"(ARG0:  $x_5$ )  
 {  $h_7 =_q h_9$  } ⟩

670 *ʔu-čubə ti luʎ'*

⟨  $h_1, e_2$  {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 $h_3$ : "\_čubə\_v\_go.inland\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),  
 $h_5$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_6$ , BODY:  $h_7$ ),  
 $h_8$ : "\_luʎ'\_n\_older-person\_rel"(ARG0:  $x_4$ )  
 {  $h_6 =_q h_8$  } ⟩

680 *ʔu-čubə-txʷ ti luʎ'*

⟨  $h_1, e_2$  {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 $h_3$ : "\_čubə\_v\_go.inland\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),  
 $h_3$ : "\_txw\_x\_ecs\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$ ),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_luʎ'\_n\_older-person\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9$  } ⟩

690 *ʔu-čubə-c ti luʎ'*

⟨  $h_1, e_2$  {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 $h_3$ : "\_čubə\_v\_go.inland\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ ),  
 $h_3$ : "\_c\_x\_altv\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD}),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_luʎ'\_n\_older-person\_rel"(ARG0:  $x_5$ )  
 {  $h_7 =_q h_9$  } ⟩

700 *ʔu-hədʔiw̄ tsi luʔ*

$\langle h_1, e_2 \{SF \text{ PROP-OR-QUES}, E.ASPECT \text{ PERFECTIVE}\},$   
 $h_3: \text{"_hədʔiw̄\_v\_inside\_rel"}(ARG0: e_2, ARG1: x_4 \{PNG.PER \text{ 3RD}, PNG.GEND \text{ FEM}\}),$   
 $h_5: \text{"_tsi\_q\_unique.F\_rel"}(ARG0: x_4, RSTR: h_6, BODY: h_7),$   
 $h_8: \text{"_luʔ\_n\_older-person\_rel"}(ARG0: x_4)$   
 $\{ h_6 =_q h_8 \} \rangle$

710 *ʔu-hədʔiw̄-tx<sup>w</sup> tsi luʔ*

$\langle h_1, e_2 \{SF \text{ PROP-OR-QUES}, E.ASPECT \text{ PERFECTIVE}\},$   
 $h_3: \text{"_hədʔiw̄\_v\_inside\_rel"}(ARG0: e_2, ARG1: x_4 \{PNG.PER \text{ 3RD}, PNG.GEND \text{ FEM}\}),$   
 $h_3: \text{"_txw\_x\_ecs\_rel"}(ARG0: e_2, ARG1: x_5),$   
 $h_6: \text{"_tsi\_q\_unique.F\_rel"}(ARG0: x_4, RSTR: h_7, BODY: h_8),$   
 $h_9: \text{"_luʔ\_n\_older-person\_rel"}(ARG0: x_4)$   
 $\{ h_7 =_q h_9 \} \rangle$

720 *ʔu-hədʔiw̄-c tsi luʔ*

$\langle h_1, e_2 \{SF \text{ PROP-OR-QUES}, E.ASPECT \text{ PERFECTIVE}\},$   
 $h_3: \text{"_hədʔiw̄\_v\_inside\_rel"}(ARG0: e_2, ARG1: x_4),$   
 $h_3: \text{"_c\_x\_altv\_rel"}(ARG0: e_2, ARG1: x_5 \{PNG.PER \text{ 3RD}, PNG.GEND \text{ FEM}\}),$   
 $h_6: \text{"_tsi\_q\_unique.F\_rel"}(ARG0: x_5, RSTR: h_7, BODY: h_8),$   
 $h_9: \text{"_luʔ\_n\_older-person\_rel"}(ARG0: x_5)$   
 $\{ h_7 =_q h_9 \} \rangle$

730 *ʔu-ʔalil-s čəd ti sup̄qs*

$\langle h_1, e_2 \{E.ASPECT \text{ PERFECTIVE}\},$   
 $h_3: \text{"_ʔalil\_v\_go.ashore\_rel"}(ARG0: e_2, ARG1: x_4 \{PNG.PER \text{ 1ST}, PNG.NUM \text{ SG}\}),$   
 $h_3: \text{"_c\_x\_altv\_rel"}(ARG0: e_2, ARG1: x_5 \{PNG.PER \text{ 3RD}\}),$   
 $h_6: \text{"_ti\_q\_unique\_rel"}(ARG0: x_5, RSTR: h_7, BODY: h_8),$   
 $h_9: \text{"_sup̄qs\_n\_hair.seal[snohomish]\_rel"}(ARG0: x_5)$   
 $\{ h_7 =_q h_9 \} \rangle$

740 *ʔu-ʔčil-s čəxʷ ʔu ti luʔ*

$\langle h_1, e_2 \{ \text{SF QUES, E.ASPECT PERFECTIVE} \},$   
 $h_3: \text{"_ʔčil\_v\_arrive\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_4 \{ \text{PNG.PER 2ND, PNG.NUM SG} \}),$   
 $h_3: \text{"_c\_x\_altv\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_5 \{ \text{PNG.PER 3RD} \}),$   
 $h_6: \text{"_ti\_q\_unique\_rel"}(\text{ARG0: } x_5, \text{RSTR: } h_7, \text{BODY: } h_8),$   
 $h_9: \text{"_luʔ\_n\_older-person\_rel"}(\text{ARG0: } x_5)$   
 $\{ h_7 =_q h_9 \} \rangle$

750 *ʔu-təlawil-s čəʔ ti sqʷəbay?*

$\langle h_1, e_2 \{ \text{E.ASPECT PERFECTIVE} \},$   
 $h_3: \text{"_təlawil\_v\_run\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_4 \{ \text{PNG.PER 1ST, PNG.NUM PL} \}),$   
 $h_3: \text{"_c\_x\_altv\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_5 \{ \text{PNG.PER 3RD} \}),$   
 $h_6: \text{"_ti\_q\_unique\_rel"}(\text{ARG0: } x_5, \text{RSTR: } h_7, \text{BODY: } h_8),$   
 $h_9: \text{"_sqʷəbay?\_n\_dog\_rel"}(\text{ARG0: } x_5)$   
 $\{ h_7 =_q h_9 \} \rangle$

760 *ʔu-ʔuʔxʷ-c čəlap ʔu tsi čəčas*

$\langle h_1, e_2 \{ \text{SF QUES, E.ASPECT PERFECTIVE} \},$   
 $h_3: \text{"_ʔuʔxʷ\_v\_go\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_4 \{ \text{PNG.PER 2ND, PNG.NUM PL} \}),$   
 $h_3: \text{"_c\_x\_altv\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_5 \{ \text{PNG.PER 3RD, PNG.GEND FEM} \}),$   
 $h_6: \text{"_tsi\_q\_unique.F\_rel"}(\text{ARG0: } x_5, \text{RSTR: } h_7, \text{BODY: } h_8),$   
 $h_9: \text{"_čəčas\_n\_child\_rel"}(\text{ARG0: } x_5)$   
 $\{ h_7 =_q h_9 \} \rangle$

770 *ʔu-čubə-c čəʔ ti sqigʷac*

$\langle h_1, e_2 \{ \text{E.ASPECT PERFECTIVE} \},$   
 $h_3: \text{"_čubə\_v\_go.inland\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_4 \{ \text{PNG.PER 1ST, PNG.NUM PL} \}),$   
 $h_3: \text{"_c\_x\_altv\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_5 \{ \text{PNG.PER 3RD} \}),$   
 $h_6: \text{"_ti\_q\_unique\_rel"}(\text{ARG0: } x_5, \text{RSTR: } h_7, \text{BODY: } h_8),$   
 $h_9: \text{"_sqigʷac\_n\_deer\_rel"}(\text{ARG0: } x_5)$   
 $\{ h_7 =_q h_9 \} \rangle$

780 *ʔu-ʔəʎ'-c čəx<sup>w</sup> ʔu tsi luʎ'*

⟨  $h_1, e_2$ {SF QUES, E.ASPECT PERFECTIVE},  
 $h_3$ :"\_ʔəʎ'\_v\_come\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ {PNG.PER 2ND, PNG.NUM SG}),  
 $h_3$ :"\_c\_x\_altv\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$ {PNG.PER 3RD, PNG.GEND FEM}),  
 $h_6$ :"\_tsi\_q\_unique.F\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ :"\_luʎ'\_n\_older-person\_rel"(ARG0:  $x_5$ )  
 {  $h_7 =_q h_9$  } ⟩

790 *ʔu-hədʔiʋ'-b tsi luʎ'*

⟨  $h_1, e_2$ {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 $h_3$ :"\_hədʔiʋ'\_v\_inside\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ ),  
 $h_3$ :"\_b\_x\_csmd\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$ {PNG.PER 3RD, PNG.GEND FEM}),  
 $h_6$ :"\_tsi\_q\_unique.F\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ :"\_luʎ'\_n\_older-person\_rel"(ARG0:  $x_5$ )  
 {  $h_7 =_q h_9$  } ⟩

800 *ʔu-hədʔiʋ'-t tsi luʎ'*

⟨  $h_1, e_2$ {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 $h_3$ :"\_hədʔiʋ'\_v\_inside\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ {PNG.PER 3RD, PNG.GEND FEM}),  
 $h_3$ :"\_t\_x\_ics\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$ ),  
 $h_6$ :"\_tsi\_q\_unique.F\_rel"(ARG0:  $x_4$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ :"\_luʎ'\_n\_older-person\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9$  } ⟩

810 *ʔu-ʔuʎ<sup>w</sup> ti čăčas*

⟨  $h_1, e_2$ {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 $h_3$ :"\_ʔuʎ<sup>w</sup>\_v\_go\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ {PNG.PER 3RD}),  
 $h_5$ :"\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_6$ , BODY:  $h_7$ ),  
 $h_8$ :"\_čăčas\_n\_child\_rel"(ARG0:  $x_4$ )  
 {  $h_6 =_q h_8$  } ⟩

820 *ʔu-ʔuχ<sup>w</sup>-tx<sup>w</sup> ti čačas*

$\langle h_1, e_2 \{SF \text{ PROP-OR-QUES}, E.ASPECT \text{ PERFECTIVE}\},$   
 $h_3: \text{"_ʔu}\chi^w\_v\_go\_rel"}(ARG0: e_2, ARG1: x_4 \{PNG.PER \text{ 3RD}\}),$   
 $h_3: \text{"_txw\_x\_ecs\_rel"}(ARG0: e_2, ARG1: x_5),$   
 $h_6: \text{"_ti\_q\_unique\_rel"}(ARG0: x_4, RSTR: h_7, BODY: h_8),$   
 $h_9: \text{"_}\acute{c}a\acute{c}as\_n\_child\_rel"}(ARG0: x_4)$   
 $\{ h_7 =_q h_9 \} \rangle$

830 *ʔu-ʔəλ<sup>w</sup>-tx<sup>w</sup> ti čačas*

$\langle h_1, e_2 \{SF \text{ PROP-OR-QUES}, E.ASPECT \text{ PERFECTIVE}\},$   
 $h_3: \text{"_ʔ}\acute{\lambda}^w\_v\_come\_rel"}(ARG0: e_2, ARG1: x_4 \{PNG.PER \text{ 3RD}\}),$   
 $h_3: \text{"_txw\_x\_ecs\_rel"}(ARG0: e_2, ARG1: x_5),$   
 $h_6: \text{"_ti\_q\_unique\_rel"}(ARG0: x_4, RSTR: h_7, BODY: h_8),$   
 $h_9: \text{"_}\acute{c}a\acute{c}as\_n\_child\_rel"}(ARG0: x_4)$   
 $\{ h_7 =_q h_9 \} \rangle$

840 *ʔu-tuk<sup>w</sup> ti čačas*

$\langle h_1, e_2 \{SF \text{ PROP-OR-QUES}, E.ASPECT \text{ PERFECTIVE}\},$   
 $h_3: \text{"_tuk}^w\_v\_go.home\_rel"}(ARG0: e_2, ARG1: x_4 \{PNG.PER \text{ 3RD}\}),$   
 $h_5: \text{"_ti\_q\_unique\_rel"}(ARG0: x_4, RSTR: h_6, BODY: h_7),$   
 $h_8: \text{"_}\acute{c}a\acute{c}as\_n\_child\_rel"}(ARG0: x_4)$   
 $\{ h_6 =_q h_8 \} \rangle$

850 *ʔu-tuk<sup>w</sup>-tx<sup>w</sup> ti čačas*

$\langle h_1, e_2 \{SF \text{ PROP-OR-QUES}, E.ASPECT \text{ PERFECTIVE}\},$   
 $h_3: \text{"_tuk}^w\_v\_go.home\_rel"}(ARG0: e_2, ARG1: x_4 \{PNG.PER \text{ 3RD}\}),$   
 $h_3: \text{"_txw\_x\_ecs\_rel"}(ARG0: e_2, ARG1: x_5),$   
 $h_6: \text{"_ti\_q\_unique\_rel"}(ARG0: x_4, RSTR: h_7, BODY: h_8),$   
 $h_9: \text{"_}\acute{c}a\acute{c}as\_n\_child\_rel"}(ARG0: x_4)$   
 $\{ h_7 =_q h_9 \} \rangle$

- 860 *ʔu-ʔəʎ'ti čačas*  
 ⟨  $h_1, e_2$ {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 |  $h_3$ :"\_ʔəʎ'\_v\_come\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ {PNG.PER 3RD}),  
 |  $h_5$ :"\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_6$ , BODY:  $h_7$ ),  
 |  $h_8$ :"\_čačas\_n\_child\_rel"(ARG0:  $x_4$ )  
 | {  $h_6 =_q h_8$  } ⟩
- 870 *ʔu-ʔuχ<sup>w</sup>*  
 ⟨  $h_1, e_2$ {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 |  $h_3$ :"\_ʔuχ<sup>w</sup>\_v\_go\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ ) |  
 | {} ⟩
- 880 *ʔu-ʔuχ<sup>w</sup>-tx<sup>w</sup>*  
 ⟨  $h_1, e_2$ {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 |  $h_3$ :"\_ʔuχ<sup>w</sup>\_v\_go\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ ),  
 |  $h_3$ :"\_txw\_x\_ecs\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$ ) |  
 | {} ⟩
- 890 *ʔu-ʔəʎ'-tx<sup>w</sup>*  
 ⟨  $h_1, e_2$ {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 |  $h_3$ :"\_ʔəʎ'\_v\_come\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ ),  
 |  $h_3$ :"\_txw\_x\_ecs\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$ ) |  
 | {} ⟩
- 900 *ʔu-ʔuk<sup>w</sup>*  
 ⟨  $h_1, e_2$ {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 |  $h_3$ :"\_ʔuk<sup>w</sup>\_v\_go.home\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ ) |  
 | {} ⟩
- 910 *ʔu-ʔuk<sup>w</sup>-tx<sup>w</sup>*  
 ⟨  $h_1, e_2$ {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 |  $h_3$ :"\_ʔuk<sup>w</sup>\_v\_go.home\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ ),  
 |  $h_3$ :"\_txw\_x\_ecs\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$ ) |  
 | {} ⟩
- 920 *ʔu-ʔəʎ'*  
 ⟨  $h_1, e_2$ {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 |  $h_3$ :"\_ʔəʎ'\_v\_come\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ ) |  
 | {} ⟩

930 *ʔu-gʷəč-b ti čačas ʔə ti sqəlalitut*

⟨  $h_1, e_2$  {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 $h_3$ : "\_gʷəč\_v\_look.for\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),  
 $h_3$ : "\_b\_x\_csmd\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD}),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_čačas\_n\_child\_rel"(ARG0:  $x_5$ ),  
 $h_{10}$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_{11}$ , BODY:  $h_{12}$ ),  
 $h_{13}$ : "\_sqəlalitut\_n\_spirit.power\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9, h_{11} =_q h_{13}$  } ⟩

940 *ʔu-čʷəl-b tsi luλ'ʔə ti sʔuladxʷ*

⟨  $h_1, e_2$  {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 $h_3$ : "\_čʷəl\_v\_cooked/ripened\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),  
 $h_3$ : "\_b\_x\_csmd\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD, PNG.GEND FEM}),  
 $h_6$ : "\_tsi\_q\_unique.F\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_luλ'\_n\_older-person\_rel"(ARG0:  $x_5$ ),  
 $h_{10}$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_{11}$ , BODY:  $h_{12}$ ),  
 $h_{13}$ : "\_sʔuladxʷ\_n\_salmon\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9, h_{11} =_q h_{13}$  } ⟩

950 *ʔu-hədʔiʷ-b ti luλ'ʔə ti hud*

⟨  $h_1, e_2$  {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 $h_3$ : "\_hədʔiʷ\_v\_inside\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),  
 $h_3$ : "\_b\_x\_csmd\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD}),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_luλ'\_n\_older-person\_rel"(ARG0:  $x_5$ ),  
 $h_{10}$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_{11}$ , BODY:  $h_{12}$ ),  
 $h_{13}$ : "\_hud\_n\_fire/heat\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9, h_{11} =_q h_{13}$  } ⟩

960 *ʔu-gʷəč̣-b čəd ʔə ti sqəlalitut*

⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},  
 $h_3$ : "\_gʷəč̣\_v\_look.for\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),  
 $h_3$ : "\_b\_x\_csmd\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 1ST, PNG.NUM SG}),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_sqəlalitut\_n\_spirit.power\_rel"(ARG0:  $x_4$ )  
{  $h_7 =_q h_9$  } ⟩

970 *ʔu-ǰʷəl-b čəd ʔə ti sʔuladxʷ*

⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},  
 $h_3$ : "\_ǰʷəl\_v\_cooked/ripened\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),  
 $h_3$ : "\_b\_x\_csmd\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 1ST, PNG.NUM SG}),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_sʔuladxʷ\_n\_salmon\_rel"(ARG0:  $x_4$ )  
{  $h_7 =_q h_9$  } ⟩

980 *ʔu-hədʔiʷ-b čəd ʔə ti hud*

⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},  
 $h_3$ : "\_hədʔiʷ\_v\_inside\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),  
 $h_3$ : "\_b\_x\_csmd\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 1ST, PNG.NUM SG}),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_hud\_n\_fire/heat\_rel"(ARG0:  $x_4$ )  
{  $h_7 =_q h_9$  } ⟩

990 *ʔu-gʷəč̣-b ti čačas ʔə tsi čačas*

⟨  $h_1, e_2$  {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 $h_3$ : "\_gʷəč̣\_v\_look.for\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD, PNG.GEND FEM}),  
 $h_3$ : "\_b\_x\_csmd\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD}),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_čačas\_n\_child\_rel"(ARG0:  $x_5$ ),  
 $h_{10}$ : "\_tsi\_q\_unique.F\_rel"(ARG0:  $x_4$ , RSTR:  $h_{11}$ , BODY:  $h_{12}$ ),  
 $h_{13}$ : "\_čačas\_n\_child\_rel"(ARG0:  $x_4$ )  
{  $h_7 =_q h_9, h_{11} =_q h_{13}$  } ⟩

1000 *ʔu-gʷəč̣-t-b ʔə ti čačas tsi čačas*

⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},

$h_3$ : "\_gʷəč̣\_v\_look.for\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD, PNG.GEND FEM}),

$h_3$ : "\_t\_x\_ics\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD}),

$h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),

$h_9$ : "\_čačas\_n\_child\_rel"(ARG0:  $x_5$ ),

$h_{10}$ : "\_tsi\_q\_unique.F\_rel"(ARG0:  $x_4$ , RSTR:  $h_{11}$ , BODY:  $h_{12}$ ),

$h_{13}$ : "\_čačas\_n\_child\_rel"(ARG0:  $x_4$ )

{  $h_7 =_q h_9, h_{11} =_q h_{13}$  }

1010 *ʔu-gʷəč̣-t-b tsi čačas ʔə ti čačas*

⟨  $h_1, e_2$  {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},

$h_3$ : "\_gʷəč̣\_v\_look.for\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD, PNG.GEND FEM}),

$h_3$ : "\_t\_x\_ics\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD}),

$h_6$ : "\_tsi\_q\_unique.F\_rel"(ARG0:  $x_4$ , RSTR:  $h_7$ , BODY:  $h_8$ ),

$h_9$ : "\_čačas\_n\_child\_rel"(ARG0:  $x_4$ ),

$h_{10}$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_{11}$ , BODY:  $h_{12}$ ),

$h_{13}$ : "\_čačas\_n\_child\_rel"(ARG0:  $x_5$ )

{  $h_7 =_q h_9, h_{11} =_q h_{13}$  }

1020 *ʔu-č̣ʷəl-b ʔə ti sʔuladxʷ*

⟨  $h_1, e_2$  {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},

$h_3$ : "\_č̣ʷəl\_v\_cooked/ripened\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),

$h_3$ : "\_b\_x\_csmd\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$ ),

$h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_7$ , BODY:  $h_8$ ),

$h_9$ : "\_sʔuladxʷ\_n\_salmon\_rel"(ARG0:  $x_4$ )

{  $h_7 =_q h_9$  }



1070 *ʔu-šab-alik<sup>w</sup> tsi luλ'*

$\langle h_1, e_2 \{SF \text{ PROP-OR-QUES, E.ASPECT PERFECTIVE}\},$   
 $h_3: \text{"_šab\_v\_dry\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_4),$   
 $h_3: \text{"_alikh\_x\_act\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_5 \{PNG.PER 3RD, PNG.GEND FEM\}),$   
 $h_6: \text{"_tsi\_q\_unique.F\_rel"}(\text{ARG0: } x_5, \text{RSTR: } h_7, \text{BODY: } h_8),$   
 $h_9: \text{"_luλ'\_n\_older-person\_rel"}(\text{ARG0: } x_5)$   
 $\{ h_7 =_q h_9 \} \rangle$

1080 *ʔu-šab-alik<sup>w</sup> tsi luλ'ʔə ti sʔuladx<sup>w</sup>*

$\langle h_1, e_2 \{SF \text{ PROP-OR-QUES, E.ASPECT PERFECTIVE}\},$   
 $h_3: \text{"_šab\_v\_dry\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_4 \{PNG.PER 3RD\}),$   
 $h_3: \text{"_alikh\_x\_act\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_5 \{PNG.PER 3RD, PNG.GEND FEM\}),$   
 $h_6: \text{"_tsi\_q\_unique.F\_rel"}(\text{ARG0: } x_5, \text{RSTR: } h_7, \text{BODY: } h_8),$   
 $h_9: \text{"_luλ'\_n\_older-person\_rel"}(\text{ARG0: } x_5),$   
 $h_{10}: \text{"_ti\_q\_unique\_rel"}(\text{ARG0: } x_4, \text{RSTR: } h_{11}, \text{BODY: } h_{12}),$   
 $h_{13}: \text{"_sʔuladx<sup>w</sup>\_n\_salmon\_rel"}(\text{ARG0: } x_4)$   
 $\{ h_7 =_q h_9, h_{11} =_q h_{13} \} \rangle$

1090 *ʔu-huyalc ti luλ'*

$\langle h_1, e_2 \{SF \text{ PROP-OR-QUES, E.ASPECT PERFECTIVE}\},$   
 $h_3: \text{"_huyalc\_v\_make/construct\_something\_rel"}(\text{ARG0: } e_2,$   
 $\text{ARG1: } x_5 \{PNG.PER 3RD\},$   
 $\text{ARG2: } x_4)$   
 $h_6: \text{"_ti\_q\_unique\_rel"}(\text{ARG0: } x_5, \text{RSTR: } h_7, \text{BODY: } h_8),$   
 $h_9: \text{"_luλ'\_n\_older-person\_rel"}(\text{ARG0: } x_5)$   
 $\{ h_7 =_q h_9 \} \rangle$

1100 *ʔu-huyalc ti luʔ'ʔə ti ʔalʔal*

$\langle h_1, e_2 \{SF \text{ PROP-OR-QUES}, E.ASPECT \text{ PERFECTIVE}\},$   
 $h_3: \text{"_huyalc\_v\_make/construct\_something\_rel"}(\text{ARG0: } e_2, \text{ ARG1: } x_5 \{PNG.PER \text{ 3RD}\}, \text{ ARG2: } x_4 \{PNG.PER \text{ 3RD}\}),$   
 $h_6: \text{"_ti\_q\_unique\_rel"}(\text{ARG0: } x_5, \text{ RSTR: } h_7, \text{ BODY: } h_8),$   
 $h_9: \text{"_luʔ'\_n\_older-person\_rel"}(\text{ARG0: } x_5),$   
 $h_{10}: \text{"_ti\_q\_unique\_rel"}(\text{ARG0: } x_4, \text{ RSTR: } h_{11}, \text{ BODY: } h_{12}),$   
 $h_{13}: \text{"_ʔalʔal\_n\_house\_rel"}(\text{ARG0: } x_4)$   
 $\{ h_7 =_q h_9, h_{11} =_q h_{13} \} \rangle$

1110 *ʔu-ʔuləʔ ti ʔaciʔtalbix<sup>w</sup>*

$\langle h_1, e_2 \{SF \text{ PROP-OR-QUES}, E.ASPECT \text{ PERFECTIVE}\},$   
 $h_3: \text{"_ʔuləʔ\_v\_gather\_something\_rel"}(\text{ARG0: } e_2, \text{ ARG1: } x_5 \{PNG.PER \text{ 3RD}\}, \text{ ARG2: } x_4),$   
 $h_6: \text{"_ti\_q\_unique\_rel"}(\text{ARG0: } x_5, \text{ RSTR: } h_7, \text{ BODY: } h_8),$   
 $h_9: \text{"_ʔaciʔtalbix<sup>w</sup>\_n\_person/indian\_rel"}(\text{ARG0: } x_5)$   
 $\{ h_7 =_q h_9 \} \rangle$

1120 *ʔu-ʔuləʔ ti ʔaciʔtalbix<sup>w</sup> ʔə ti bəsq<sup>w</sup>*

$\langle h_1, e_2 \{SF \text{ PROP-OR-QUES}, E.ASPECT \text{ PERFECTIVE}\},$   
 $h_3: \text{"_ʔuləʔ\_v\_gather\_something\_rel"}(\text{ARG0: } e_2, \text{ ARG1: } x_5 \{PNG.PER \text{ 3RD}\}, \text{ ARG2: } x_4 \{PNG.PER \text{ 3RD}\}),$   
 $h_6: \text{"_ti\_q\_unique\_rel"}(\text{ARG0: } x_5, \text{ RSTR: } h_7, \text{ BODY: } h_8),$   
 $h_9: \text{"_ʔaciʔtalbix<sup>w</sup>\_n\_person/indian\_rel"}(\text{ARG0: } x_5),$   
 $h_{10}: \text{"_ti\_q\_unique\_rel"}(\text{ARG0: } x_4, \text{ RSTR: } h_{11}, \text{ BODY: } h_{12}),$   
 $h_{13}: \text{"_bəsq<sup>w</sup>\_n\_crab\_rel"}(\text{ARG0: } x_4)$   
 $\{ h_7 =_q h_9, h_{11} =_q h_{13} \} \rangle$

1130 *ʔu-ʔəʔəd tsi čačas*

⟨  $h_1, e_2$  {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},

$h_3$ : "\_ʔəʔəd\_v\_eat\_something\_rel"(ARG0:  $e_2$ ,  
ARG1:  $x_5$  {PNG.PER 3RD, PNG.GEND FEM},  
ARG2:  $x_4$ )

$h_6$ : "\_tsi\_q\_unique.F\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),

$h_9$ : "\_čačas\_n\_child\_rel"(ARG0:  $x_5$ )  
{  $h_7 =_q h_9$  } )

1140 *ʔu-ʔəʔəd tsi čačas ʔə ti bəsqʷ*

⟨  $h_1, e_2$  {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},

$h_3$ : "\_ʔəʔəd\_v\_eat\_something\_rel"(ARG0:  $e_2$ ,  
ARG1:  $x_5$  {PNG.PER 3RD, PNG.GEND FEM},  
ARG2:  $x_4$  {PNG.PER 3RD})

$h_6$ : "\_tsi\_q\_unique.F\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),

$h_9$ : "\_čačas\_n\_child\_rel"(ARG0:  $x_5$ ),

$h_{10}$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_{11}$ , BODY:  $h_{12}$ ),

$h_{13}$ : "\_bəsqʷ\_n\_crab\_rel"(ARG0:  $x_4$ )  
{  $h_7 =_q h_9, h_{11} =_q h_{13}$  } )

1150 *ʔu-šab-t-b ʔə tsi luł' ti sʔuladxʷ*

⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},

$h_3$ : "\_šab\_v\_dry\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),

$h_3$ : "\_t\_x\_ics\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD, PNG.GEND FEM}),

$h_6$ : "\_tsi\_q\_unique.F\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),

$h_9$ : "\_luł'n\_older-person\_rel"(ARG0:  $x_5$ ),

$h_{10}$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_{11}$ , BODY:  $h_{12}$ ),

$h_{13}$ : "\_sʔuladxʷ\_n\_salmon\_rel"(ARG0:  $x_4$ )  
{  $h_7 =_q h_9, h_{11} =_q h_{13}$  } )

1160 *ʔu-ʔuləḥ-t-b ʔə ti luḥ' ti hud*

⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},

$h_3$ : "\_ʔuləḥ\_v\_gather\_something\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD}, ARG2:  $x_4$ ),

$h_3$ : "\_t\_x\_ics\_rel"(ARG0:  $e_2$ , ARG1:  $x_6$  {PNG.PER 3RD}),

$h_7$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_6$ , RSTR:  $h_8$ , BODY:  $h_9$ ),

$h_{10}$ : "\_luḥ' n\_older-person\_rel"(ARG0:  $x_6$ ),

$h_{11}$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_{12}$ , BODY:  $h_{13}$ ),

$h_{14}$ : "\_hud\_n\_fire/heat\_rel"(ARG0:  $x_5$ )

{  $h_8 =_q h_{10}$ ,  $h_{12} =_q h_{14}$  }

1170 *ʔu-ʔuləḥ ti luḥ' ʔə ti bəsq<sup>w</sup>*

⟨  $h_1, e_2$  {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},

$h_3$ : "\_ʔuləḥ\_v\_gather\_something\_rel"(ARG0:  $e_2$ ,  
ARG1:  $x_5$  {PNG.PER 3RD},  
ARG2:  $x_4$  {PNG.PER 3RD}),

$h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),

$h_9$ : "\_luḥ' n\_older-person\_rel"(ARG0:  $x_5$ ),

$h_{10}$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_{11}$ , BODY:  $h_{12}$ ),

$h_{13}$ : "\_bəsq<sup>w</sup>\_n\_crab\_rel"(ARG0:  $x_4$ )

{  $h_7 =_q h_9$ ,  $h_{11} =_q h_{13}$  }

1180 *ʔu-čəʔ-t-b ʔə tsi luḥ' ti sk<sup>wi</sup>?x<sup>w</sup>*

⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},

$h_3$ : "\_čəʔ\_v\_dig\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),

$h_3$ : "\_t\_x\_ics\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD, PNG.GEND FEM}),

$h_6$ : "\_tsi\_q\_unique.F\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),

$h_9$ : "\_luḥ' n\_older-person\_rel"(ARG0:  $x_5$ ),

$h_{10}$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_{11}$ , BODY:  $h_{12}$ ),

$h_{13}$ : "\_sk<sup>wi</sup>?x<sup>w</sup>\_n\_bracken.fern.rhizome\_rel"(ARG0:  $x_4$ )

{  $h_7 =_q h_9$ ,  $h_{11} =_q h_{13}$  }

1190 *ʔu-čáʔ-b tsi luʔʔə ti skʷiʔxʷ*

⟨  $h_1, e_2$  {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 $h_3$ : "\_čáʔ\_v\_dig\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),  
 $h_3$ : "\_b\_x\_csmd\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD, PNG.GEND FEM}),  
 $h_6$ : "\_tsi\_q\_unique.F\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_luʔʔ\_n\_older-person\_rel"(ARG0:  $x_5$ ),  
 $h_{10}$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_{11}$ , BODY:  $h_{12}$ ),  
 $h_{13}$ : "\_skʷiʔxʷ\_n\_bracken.fern.rhizome\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9, h_{11} =_q h_{13}$  } ⟩

1200 *ʔu-pusu-t-b ʔə ti čáčas ti sqʷəbayʔ*

⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},  
 $h_3$ : "\_pusu\_v\_be.thrown.at\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),  
 $h_3$ : "\_t\_x\_ics\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD}),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_čáčas\_n\_child\_rel"(ARG0:  $x_5$ ),  
 $h_{10}$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_{11}$ , BODY:  $h_{12}$ ),  
 $h_{13}$ : "\_sqʷəbayʔ\_n\_dog\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9, h_{11} =_q h_{13}$  } ⟩

1210 *ʔu-túču-t-b ʔə ti luʔʔə ti sčətxʷəd*

⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},  
 $h_3$ : "\_túču\_v\_be.shot.at\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),  
 $h_3$ : "\_t\_x\_ics\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD}),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_luʔʔ\_n\_older-person\_rel"(ARG0:  $x_5$ ),  
 $h_{10}$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_{11}$ , BODY:  $h_{12}$ ),  
 $h_{13}$ : "\_sčətxʷəd\_n\_bear\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9, h_{11} =_q h_{13}$  } ⟩

1220 *ʔu-túćil ti luł'ʔə ti t̥səd*

⟨  $h_1, e_2$  {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 $h_3$ : "\_túćil\_v\_shoot\_something\_rel"(ARG0:  $e_2$ ,  
 ARG1:  $x_5$  {PNG.PER 3RD},  
 ARG2:  $x_4$  {PNG.PER 3RD}),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_luł'\_n\_older-person\_rel"(ARG0:  $x_5$ ),  
 $h_{10}$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_{11}$ , BODY:  $h_{12}$ ),  
 $h_{13}$ : "\_t̥səd\_n\_arrow\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9, h_{11} =_q h_{13}$  } ⟩

1230 *ʔu-ləḵ<sup>w</sup>-t-b ʔə tsi čáčas ti bəsq<sup>w</sup>*

⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},  
 $h_3$ : "\_ləḵ<sup>w</sup>\_v\_be.eaten\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),  
 $h_3$ : "\_t\_x\_ics\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD, PNG.GEND FEM}),  
 $h_6$ : "\_tsi\_q\_unique.F\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_čáčas\_n\_child\_rel"(ARG0:  $x_5$ ),  
 $h_{10}$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_{11}$ , BODY:  $h_{12}$ ),  
 $h_{13}$ : "\_bəsq<sup>w</sup>\_n\_crab\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9, h_{11} =_q h_{13}$  } ⟩

1240 *ʔu-šab-t-b ʔə tsi luł' ti sʔuladx<sup>w</sup>*

⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},  
 $h_3$ : "\_šab\_v\_dry\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),  
 $h_3$ : "\_t\_x\_ics\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD, PNG.GEND FEM}),  
 $h_6$ : "\_tsi\_q\_unique.F\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_luł'\_n\_older-person\_rel"(ARG0:  $x_5$ ),  
 $h_{10}$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_{11}$ , BODY:  $h_{12}$ ),  
 $h_{13}$ : "\_sʔuladx<sup>w</sup>\_n\_salmon\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9, h_{11} =_q h_{13}$  } ⟩

1250 *ʔu-g<sup>w</sup>əč<sup>̣</sup>-t-b ʔə ti ʔaciɬtalbix<sup>w</sup> ti sqəlalitut*

⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},  
 $h_3$ : "\_g<sup>w</sup>əč<sup>̣</sup>\_v\_look.for\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),  
 $h_3$ : "\_t\_x\_ics\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD}),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_ʔaciɬtalbix<sup>w</sup>\_n\_person/indian\_rel"(ARG0:  $x_5$ ),  
 $h_{10}$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_{11}$ , BODY:  $h_{12}$ ),  
 $h_{13}$ : "\_sqəlalitut\_n\_spirit.power\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9, h_{11} =_q h_{13}$  } ⟩

1260 *ʔu-lək<sup>w</sup>-t-b ʔə ti sq<sup>w</sup>əbayʔ ti bəsq<sup>w</sup>*

⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},  
 $h_3$ : "\_lək<sup>w</sup>\_v\_be.eaten\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),  
 $h_3$ : "\_t\_x\_ics\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD}),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_sq<sup>w</sup>əbayʔ\_n\_dog\_rel"(ARG0:  $x_5$ ),  
 $h_{10}$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_{11}$ , BODY:  $h_{12}$ ),  
 $h_{13}$ : "\_bəsq<sup>w</sup>\_n\_crab\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9, h_{11} =_q h_{13}$  } ⟩

1270 *ʔu-ʔuləχ<sup>̣</sup> ti luχ<sup>̣</sup> ʔə ti hud*

⟨  $h_1, e_2$  {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 $h_3$ : "\_ʔuləχ<sup>̣</sup>\_v\_gather\_something\_rel"(ARG0:  $e_2$ ,  
 ARG1:  $x_5$  {PNG.PER 3RD},  
 ARG2:  $x_4$  {PNG.PER 3RD}),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_luχ<sup>̣</sup>\_n\_older-person\_rel"(ARG0:  $x_5$ ),  
 $h_{10}$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_{11}$ , BODY:  $h_{12}$ ),  
 $h_{13}$ : "\_hud\_n\_fire/heat\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9, h_{11} =_q h_{13}$  } ⟩

1280 *ʔu-čaʔ-b tsi čačas ʔə ti skwiʔx<sup>w</sup>*

⟨  $h_1, e_2$  {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 $h_3$ : "\_čaʔ\_v\_dig\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),  
 $h_3$ : "\_b\_x\_csmd\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD, PNG.GEND FEM}),  
 $h_6$ : "\_tsi\_q\_unique.F\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_čačas\_n\_child\_rel"(ARG0:  $x_5$ ),  
 $h_{10}$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_{11}$ , BODY:  $h_{12}$ ),  
 $h_{13}$ : "\_skwiʔx<sup>w</sup>\_n\_bracken.fern.rhizome\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9, h_{11} =_q h_{13}$  } ⟩

1290 *ʔu-ʔəʔəd ti čačas ʔə ti bəsq<sup>w</sup>*

⟨  $h_1, e_2$  {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 $h_3$ : "\_ʔəʔəd\_v\_eat\_something\_rel"(ARG0:  $e_2$ ,  
 ARG1:  $x_5$  {PNG.PER 3RD},  
 ARG2:  $x_4$  {PNG.PER 3RD}),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_čačas\_n\_child\_rel"(ARG0:  $x_5$ ),  
 $h_{10}$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_{11}$ , BODY:  $h_{12}$ ),  
 $h_{13}$ : "\_bəsq<sup>w</sup>\_n\_crab\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9, h_{11} =_q h_{13}$  } ⟩

1300 *ʔu-čəl-dx<sup>w</sup>-b ti sčətx<sup>w</sup>əd ʔə ti čičičič*

⟨  $h_1, e_2$  {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 $h_3$ : "\_čəl\_v\_defeated\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),  
 $h_3$ : "\_dxw\_x\_dc\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD}),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_sčətx<sup>w</sup>əd\_n\_bear\_rel"(ARG0:  $x_4$ ),  
 $h_{10}$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_{11}$ , BODY:  $h_{12}$ ),  
 $h_{13}$ : "\_čičičič<sup>w</sup>\_n\_osprey\_rel"(ARG0:  $x_5$ )  
 {  $h_7 =_q h_9, h_{11} =_q h_{13}$  } ⟩

1310 *ʔu-yiǰi-b tsi sʔadəyʔ ʔə ti spčũʔ*

$\langle h_1, e_2 \{SF \text{ PROP-OR-QUES, E.ASPECT PERFECTIVE}\},$   
 $h_3: \text{"_yiǰi\_v\_woven\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_4 \{PNG.PER 3RD\}),$   
 $h_3: \text{"_b\_x\_csm̄d\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_5 \{PNG.PER 3RD, PNG.GEND FEM\}),$   
 $h_6: \text{"_tsi\_q\_unique.F\_rel"}(\text{ARG0: } x_5, \text{RSTR: } h_7, \text{BODY: } h_8),$   
 $h_9: \text{"_sʔadəy\_n\_woman\_rel"}(\text{ARG0: } x_5),$   
 $h_{10}: \text{"_ti\_q\_unique\_rel"}(\text{ARG0: } x_4, \text{RSTR: } h_{11}, \text{BODY: } h_{12}),$   
 $h_{13}: \text{"_spčũʔ\_n\_basket\_rel"}(\text{ARG0: } x_4)$   
 $\{ h_7 =_q h_9, h_{11} =_q h_{13} \} \rangle$

1320 *ʔu-yiǰi-t-b tsi sʔadəyʔ ʔə ti spčũʔ*

$\langle h_1, e_2 \{SF \text{ PROP-OR-QUES, E.ASPECT PERFECTIVE}\},$   
 $h_3: \text{"_yiǰi\_v\_woven\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_4 \{PNG.PER 3RD, PNG.GEND FEM\}),$   
 $h_3: \text{"_t\_x\_ics\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_5 \{PNG.PER 3RD\}),$   
 $h_6: \text{"_tsi\_q\_unique.F\_rel"}(\text{ARG0: } x_4, \text{RSTR: } h_7, \text{BODY: } h_8),$   
 $h_9: \text{"_sʔadəy\_n\_woman\_rel"}(\text{ARG0: } x_4),$   
 $h_{10}: \text{"_ti\_q\_unique\_rel"}(\text{ARG0: } x_5, \text{RSTR: } h_{11}, \text{BODY: } h_{12}),$   
 $h_{13}: \text{"_spčũʔ\_n\_basket\_rel"}(\text{ARG0: } x_5)$   
 $\{ h_7 =_q h_9, h_{11} =_q h_{13} \} \rangle$

1330 *ʔu-čəl-alik<sup>w</sup> tsi ʔaʔačapəd*

$\langle h_1, e_2 \{SF \text{ PROP-OR-QUES, E.ASPECT PERFECTIVE}\},$   
 $h_3: \text{"_čəl\_v\_defeated\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_4),$   
 $h_3: \text{"_alik<sup>w</sup>\_x\_act\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_5 \{PNG.PER 3RD, PNG.GEND FEM\}),$   
 $h_6: \text{"_tsi\_q\_unique.F\_rel"}(\text{ARG0: } x_5, \text{RSTR: } h_7, \text{BODY: } h_8),$   
 $h_9: \text{"_ʔaʔačapəd\_n\_ant\_rel"}(\text{ARG0: } x_5)$   
 $\{ h_7 =_q h_9 \} \rangle$

1340 *ʔu-yiǰi-t ti spčũ?*

$\langle h_1, e_2 \{ \text{SF PROP-OR-QUES, E.ASPECT PERFECTIVE} \},$   
 $h_3: \text{"_yiǰi\_v\_woven\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_4 \{ \text{PNG.PER 3RD} \}),$   
 $h_3: \text{"_t\_x\_ics\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_5),$   
 $h_6: \text{"_ti\_q\_unique\_rel"}(\text{ARG0: } x_4, \text{RSTR: } h_7, \text{BODY: } h_8),$   
 $h_9: \text{"_spčũ?\_n\_basket\_rel"}(\text{ARG0: } x_4)$   
 $\{ h_7 =_q h_9 \} \rangle$

1350 *ʔu-pusil ti lægʷəb*

$\langle h_1, e_2 \{ \text{SF PROP-OR-QUES, E.ASPECT PERFECTIVE} \},$   
 $h_3: \text{"_pusil\_v\_throw\_something\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_5 \{ \text{PNG.PER 3RD} \}, \text{ARG2: } x_4),$   
 $h_6: \text{"_ti\_q\_unique\_rel"}(\text{ARG0: } x_5, \text{RSTR: } h_7, \text{BODY: } h_8),$   
 $h_9: \text{"_lægʷəb\_n\_young.person\_rel"}(\text{ARG0: } x_5)$   
 $\{ h_7 =_q h_9 \} \rangle$

1360 *ʔu-yəc-b ti stubš ʔə ti syəcəb*

$\langle h_1, e_2 \{ \text{SF PROP-OR-QUES, E.ASPECT PERFECTIVE} \},$   
 $h_3: \text{"_yəc\_v\_be.reported\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_4 \{ \text{PNG.PER 3RD} \}),$   
 $h_3: \text{"_b\_x\_csm�\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_5 \{ \text{PNG.PER 3RD} \}),$   
 $h_6: \text{"_ti\_q\_unique\_rel"}(\text{ARG0: } x_5, \text{RSTR: } h_7, \text{BODY: } h_8),$   
 $h_9: \text{"_stubš\_n\_man\_rel"}(\text{ARG0: } x_5),$   
 $h_{10}: \text{"_ti\_q\_unique\_rel"}(\text{ARG0: } x_4, \text{RSTR: } h_{11}, \text{BODY: } h_{12}),$   
 $h_{13}: \text{"_syəcəb\_n\_news\_rel"}(\text{ARG0: } x_4)$   
 $\{ h_7 =_q h_9, h_{11} =_q h_{13} \} \rangle$

1370 *ʔu-ʔiči-b ʔə ti ʔulal*

$\langle h_1, e_2 \{ \text{SF PROP-OR-QUES, E.ASPECT PERFECTIVE} \},$   
 $h_3: \text{"_ʔiči\_v\_be.cut.with.knife\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_4 \{ \text{PNG.PER 3RD} \}),$   
 $h_3: \text{"_b\_x\_csm�\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_5),$   
 $h_6: \text{"_ti\_q\_unique\_rel"}(\text{ARG0: } x_4, \text{RSTR: } h_7, \text{BODY: } h_8),$   
 $h_9: \text{"_ʔulal\_n\_cattail\_rel"}(\text{ARG0: } x_4)$   
 $\{ h_7 =_q h_9 \} \rangle$

1380 *ʔu-pusu-t-b ʔə ti čáčas ti sbiaw*

$\langle h_1, e_2 \{E.ASPECT PERFECTIVE\},$   
 $h_3: \text{"_pusu\_v\_be.thrown.at\_rel"}(ARG0: e_2, ARG1: x_4 \{PNG.PER 3RD\}),$   
 $h_3: \text{"_t\_x\_ics\_rel"}(ARG0: e_2, ARG1: x_5 \{PNG.PER 3RD\}),$   
 $h_6: \text{"_ti\_q\_unique\_rel"}(ARG0: x_5, RSTR: h_7, BODY: h_8),$   
 $h_9: \text{"_čáčas\_n\_child\_rel"}(ARG0: x_5),$   
 $h_{10}: \text{"_ti\_q\_unique\_rel"}(ARG0: x_4, RSTR: h_{11}, BODY: h_{12}),$   
 $h_{13}: \text{"_sbiaw\_n\_coyote\_rel"}(ARG0: x_4)$   
 $\{ h_7 =_q h_9, h_{11} =_q h_{13} \} \rangle$

1390 *ʔu-pusil ʔə ti čłá?*

$\langle h_1, e_2 \{SF PROP-OR-QUES, E.ASPECT PERFECTIVE\},$   
 $h_3: \text{"_pusil\_v\_throw\_something\_rel"}(ARG0: e_2, ARG1: x_5, ARG2: x_4 \{PNG.PER 3RD\}),$   
 $h_6: \text{"_ti\_q\_unique\_rel"}(ARG0: x_4, RSTR: h_7, BODY: h_8),$   
 $h_9: \text{"_čłá?_n\_rock\_rel"}(ARG0: x_4)$   
 $\{ h_7 =_q h_9 \} \rangle$

1400 *ʔu-ʔəʔəd ti q<sup>w</sup>ist ʔə ti sq<sup>w</sup>i?q<sup>w</sup>ali?*

$\langle h_1, e_2 \{SF PROP-OR-QUES, E.ASPECT PERFECTIVE\},$   
 $h_3: \text{"_ʔəʔəd\_v\_eat\_something\_rel"}(ARG0: e_2,$   
 $ARG1: x_5 \{PNG.PER 3RD\}$   
 $ARG2: x_4 \{PNG.PER 3RD\})$   
 $h_6: \text{"_ti\_q\_unique\_rel"}(ARG0: x_5, RSTR: h_7, BODY: h_8),$   
 $h_9: \text{"_q<sup>w</sup>ist\_n\_bovine\_rel"}(ARG0: x_5),$   
 $h_{10}: \text{"_ti\_q\_unique\_rel"}(ARG0: x_4, RSTR: h_{11}, BODY: h_{12}),$   
 $h_{13}: \text{"_sq<sup>w</sup>i?q<sup>w</sup>ali?_n\_hay\_rel"}(ARG0: x_4)$   
 $\{ h_7 =_q h_9, h_{11} =_q h_{13} \} \rangle$

- 1410 *ʔu-ʔusil-s-b ʔə ti čičičič ti sʔuladx<sup>w</sup>*  
 ⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},  
 $h_3$ : "\_ʔusil\_v\_dive\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),  
 $h_3$ : "\_c\_x\_altv\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD}),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_čičičič<sup>w</sup>\_n\_osprey\_rel"(ARG0:  $x_4$ ),  
 $h_{10}$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_{11}$ , BODY:  $h_{12}$ ),  
 $h_{13}$ : "\_sʔuladx<sup>w</sup>\_n\_salmon\_rel"(ARG0:  $x_5$ )  
 {  $h_7 =_q h_9, h_{11} =_q h_{13}$  } ⟩
- 1420 *ʔu-ʔuḫ<sup>w</sup> ti luḫ'*  
 ⟨  $h_1, e_2$  {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 $h_3$ : "\_ʔuḫ<sup>w</sup>\_v\_go\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),  
 $h_5$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_6$ , BODY:  $h_7$ ),  
 $h_8$ : "\_luḫ'<sup>w</sup>\_n\_older-person\_rel"(ARG0:  $x_4$ )  
 {  $h_6 =_q h_8$  } ⟩
- 1430 *ʔu-ʔuḫ<sup>w</sup>-c ti čáčas*  
 ⟨  $h_1, e_2$  {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 $h_3$ : "\_ʔuḫ<sup>w</sup>\_v\_go\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ ),  
 $h_3$ : "\_c\_x\_altv\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD}),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_čáčas\_n\_child\_rel"(ARG0:  $x_5$ )  
 {  $h_7 =_q h_9$  } ⟩
- 1440 *ʔu-ʔuḫ<sup>w</sup>-c-b ʔə ti luḫ' ti čáčas*  
 ⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},  
 $h_3$ : "\_ʔuḫ<sup>w</sup>\_v\_go\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),  
 $h_3$ : "\_c\_x\_altv\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD}),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_luḫ'<sup>w</sup>\_n\_older-person\_rel"(ARG0:  $x_4$ ),  
 $h_{10}$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_{11}$ , BODY:  $h_{12}$ ),  
 $h_{13}$ : "\_čáčas\_n\_child\_rel"(ARG0:  $x_5$ )  
 {  $h_7 =_q h_9, h_{11} =_q h_{13}$  } ⟩

1450 *ʔu-ʔuχ<sup>w</sup>-tx<sup>w</sup> ti čačas*

⟨  $h_1, e_2$  {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 $h_3$ : "\_ʔuχ<sup>w</sup>\_v\_go\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),  
 $h_3$ : "\_txw\_x\_ecs\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$ ),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_čačas\_n\_child\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9$  } ⟩

1460 *ʔu-ʔuχ<sup>w</sup>-tx<sup>w</sup>-b ʔə ti luχ' ti čačas*

⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},  
 $h_3$ : "\_ʔuχ<sup>w</sup>\_v\_go\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),  
 $h_3$ : "\_txw\_x\_ecs\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD}),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_luχ'\_n\_older-person\_rel"(ARG0:  $x_5$ ),  
 $h_{10}$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_{11}$ , BODY:  $h_{12}$ ),  
 $h_{13}$ : "\_čačas\_n\_child\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9, h_{11} =_q h_{13}$  } ⟩

1470 *ʔu-ʔəỵ-dx<sup>w</sup>-b ʔə ti čačas ti sqwəbay?*

⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},  
 $h_3$ : "\_ʔəỵ\_v\_found\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),  
 $h_3$ : "\_dxw\_x\_dc\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD}),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_čačas\_n\_child\_rel"(ARG0:  $x_5$ ),  
 $h_{10}$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_{11}$ , BODY:  $h_{12}$ ),  
 $h_{13}$ : "\_sqwəbay?\_n\_dog\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9, h_{11} =_q h_{13}$  } ⟩

1480 *ʔu-ʔəy̌-dx<sup>w</sup>-b ʔə ti sq<sup>w</sup>əbay? ti čáčas*

⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},  
 $h_3$ : "\_ʔəy̌\_v\_found\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),  
 $h_3$ : "\_dxw\_x\_dc\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD}),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_sq<sup>w</sup>əbay?\_n\_dog\_rel"(ARG0:  $x_5$ ),  
 $h_{10}$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_{11}$ , BODY:  $h_{12}$ ),  
 $h_{13}$ : "\_čáčas\_n\_child\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9, h_{11} =_q h_{13}$  } ⟩

1490 *ʔu-k<sup>w</sup>əda-t-b ʔə tsi čáčas ti sq<sup>w</sup>əbay?*

⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},  
 $h_3$ : "\_k<sup>w</sup>əda\_v\_take/grab\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),  
 $h_3$ : "\_t\_x\_ics\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD, PNG.GEND FEM}),  
 $h_6$ : "\_tsi\_q\_unique.F\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_čáčas\_n\_child\_rel"(ARG0:  $x_5$ ),  
 $h_{10}$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_{11}$ , BODY:  $h_{12}$ ),  
 $h_{13}$ : "\_sq<sup>w</sup>əbay?\_n\_dog\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9, h_{11} =_q h_{13}$  } ⟩

1500 *ʔu-ʔusil-s-b ti sʔuladx<sup>w</sup>*

⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},  
 $h_3$ : "\_ʔusil\_v\_dive\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ ),  
 $h_3$ : "\_c\_x\_altv\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD}),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_sʔuladx<sup>w</sup>\_n\_salmon\_rel"(ARG0:  $x_5$ )  
 {  $h_7 =_q h_9$  } ⟩

1510 *ʔu-ʔusil-s-b ʔə ti ćiḥćiḥ*

$\langle h_1, e_2 \{E.ASPECT PERFECTIVE\},$   
 $h_3: \_?usil\_v\_dive\_rel(ARG0: e_2, ARG1: x_4 \{PNG.PER 3RD\}),$   
 $h_3: \_c\_x\_altv\_rel(ARG0: e_2, ARG1: i_5),$   
 $h_6: \_ti\_q\_unique\_rel(ARG0: x_4, RSTR: h_7, BODY: h_8),$   
 $h_9: \_ćiḥćiḥ^w\_n\_osprey\_rel(ARG0: x_4)$   
 $\{ h_7 =_q h_9 \} \rangle$

1520 *ʔu-ʔusil-s-b*

$\langle h_1, e_2 \{E.ASPECT PERFECTIVE\},$   
 $h_3: \_?usil\_v\_dive\_rel(ARG0: e_2, ARG1: x_4),$   
 $h_3: \_c\_x\_altv\_rel(ARG0: e_2, ARG1: i_5)$   
 $\{ \} \rangle$

1530 *ʔu-kʷəda-t-b ti sqʷəbay?*

$\langle h_1, e_2 \{E.ASPECT PERFECTIVE\},$   
 $h_3: \_kʷəda\_v\_take/grab\_rel(ARG0: e_2, ARG1: x_4 \{PNG.PER 3RD\}),$   
 $h_3: \_t\_x\_ics\_rel(ARG0: e_2, ARG1: i_5),$   
 $h_6: \_ti\_q\_unique\_rel(ARG0: x_4, RSTR: h_7, BODY: h_8),$   
 $h_9: \_sqʷəbay?_n\_dog\_rel(ARG0: x_4)$   
 $\{ h_7 =_q h_9 \} \rangle$

1540 *ʔu-kʷəda-t-b ʔə tsi ćačas*

$\langle h_1, e_2 \{E.ASPECT PERFECTIVE\},$   
 $h_3: \_kʷəda\_v\_take/grab\_rel(ARG0: e_2, ARG1: x_4),$   
 $h_3: \_t\_x\_ics\_rel(ARG0: e_2, ARG1: x_5 \{PNG.PER 3RD, PNG.GEND FEM\}),$   
 $h_6: \_tsi\_q\_unique.F\_rel(ARG0: x_5, RSTR: h_7, BODY: h_8),$   
 $h_9: \_ćačas\_n\_child\_rel(ARG0: x_5)$   
 $\{ h_7 =_q h_9 \} \rangle$

1550 *ʔu-kʷəda-t-b*

$\langle h_1, e_2 \{E.ASPECT PERFECTIVE\},$   
 $h_3: \_kʷəda\_v\_take/grab\_rel(ARG0: e_2, ARG1: x_4),$   
 $h_3: \_t\_x\_ics\_rel(ARG0: e_2, ARG1: i_5)$   
 $\{ \} \rangle$

1560 *ʔu-gʷəč̣-t-b ʔə ti č̣ač̣as ti sqʷəbay?*

⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},

$h_3$ : "\_gʷəč̣\_v\_look.for\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),

$h_3$ : "\_t\_x\_ics\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD}),

$h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),

$h_9$ : "\_č̣ač̣as\_n\_child\_rel"(ARG0:  $x_5$ ),

$h_{10}$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_{11}$ , BODY:  $h_{12}$ ),

$h_{13}$ : "\_sqʷəbay?\_n\_dog\_rel"(ARG0:  $x_4$ )

{  $h_7 =_q h_9, h_{11} =_q h_{13}$  }

1570 *ʔu-gʷəč̣-t-b ti sqʷəbay? ʔə ti č̣ač̣as*

⟨  $h_1, e_2$  {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},

$h_3$ : "\_gʷəč̣\_v\_look.for\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),

$h_3$ : "\_t\_x\_ics\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD}),

$h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_7$ , BODY:  $h_8$ ),

$h_9$ : "\_sqʷəbay?\_n\_dog\_rel"(ARG0:  $x_4$ ),

$h_{10}$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_{11}$ , BODY:  $h_{12}$ ),

$h_{13}$ : "\_č̣ač̣as\_n\_child\_rel"(ARG0:  $x_5$ )

{  $h_7 =_q h_9, h_{11} =_q h_{13}$  }

1580 *ʔu-gʷəč̣-t-b ʔə ti sqʷəbay? ti č̣ač̣as*

⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},

$h_3$ : "\_gʷəč̣\_v\_look.for\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),

$h_3$ : "\_t\_x\_ics\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD}),

$h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),

$h_9$ : "\_sqʷəbay?\_n\_dog\_rel"(ARG0:  $x_5$ ),

$h_{10}$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_{11}$ , BODY:  $h_{12}$ ),

$h_{13}$ : "\_č̣ač̣as\_n\_child\_rel"(ARG0:  $x_4$ )

{  $h_7 =_q h_9, h_{11} =_q h_{13}$  }

1590 *ʔu-g<sup>w</sup>əč-t-b ti čačas ʔə ti sq<sup>w</sup>əbay?*

⟨  $h_1, e_2$  {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 $h_3$ : "\_g<sup>w</sup>əč\_v\_look.for\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),  
 $h_3$ : "\_t\_x\_ics\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD}),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_čačas\_n\_child\_rel"(ARG0:  $x_4$ ),  
 $h_{10}$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_{11}$ , BODY:  $h_{12}$ ),  
 $h_{13}$ : "\_sq<sup>w</sup>əbay?\_n\_dog\_rel"(ARG0:  $x_5$ )  
 {  $h_7 =_q h_9, h_{11} =_q h_{13}$  } ⟩

1600 *ʔu-túk<sup>w</sup>-tx<sup>w</sup>-b ʔə ti čačas tsi čačas*

⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},  
 $h_3$ : "\_túk<sup>w</sup>\_v\_go.home\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD, PNG.GEND FEM}),  
 $h_3$ : "\_txw\_x\_ecs\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD}),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_čačas\_n\_child\_rel"(ARG0:  $x_5$ ),  
 $h_{10}$ : "\_tsi\_q\_unique.F\_rel"(ARG0:  $x_4$ , RSTR:  $h_{11}$ , BODY:  $h_{12}$ ),  
 $h_{13}$ : "\_čačas\_n\_child\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9, h_{11} =_q h_{13}$  } ⟩

1610 *ʔu-túk<sup>w</sup>-tx<sup>w</sup>-b tsi čačas ʔə ti čačas*

⟨  $h_1, e_2$  {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 $h_3$ : "\_túk<sup>w</sup>\_v\_go.home\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD, PNG.GEND FEM}),  
 $h_3$ : "\_txw\_x\_ecs\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD}),  
 $h_6$ : "\_tsi\_q\_unique.F\_rel"(ARG0:  $x_4$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_čačas\_n\_child\_rel"(ARG0:  $x_4$ ),  
 $h_{10}$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_{11}$ , BODY:  $h_{12}$ ),  
 $h_{13}$ : "\_čačas\_n\_child\_rel"(ARG0:  $x_5$ )  
 {  $h_7 =_q h_9, h_{11} =_q h_{13}$  } ⟩

1620 *ʔu-túk<sup>w</sup>-tx<sup>w</sup>-b ʔə tsi čáčas ti čáčas*

$\langle h_1, e_2 \{E.ASPECT PERFECTIVE\},$

$h_3: \text{"_túk<sup>w</sup>_v_go.home_rel"}(ARG0: e_2, ARG1: x_4 \{PNG.PER 3RD\}),$

$h_3: \text{"_txw_x_ecs_rel"}(ARG0: e_2, ARG1: x_5 \{PNG.PER 3RD, PNG.GEND FEM\}),$

$h_6: \text{"_tsi_q_unique.F_rel"}(ARG0: x_5, RSTR: h_7, BODY: h_8),$

$h_9: \text{"_čáčas_n_child_rel"}(ARG0: x_5),$

$h_{10}: \text{"_ti_q_unique_rel"}(ARG0: x_4, RSTR: h_{11}, BODY: h_{12}),$

$h_{13}: \text{"_čáčas_n_child_rel"}(ARG0: x_4)$

$\{ h_7 =_q h_9, h_{11} =_q h_{13} \} \rangle$

1630 *ʔu-túk<sup>w</sup>-tx<sup>w</sup>-b ti čáčas ʔə tsi čáčas*

$\langle h_1, e_2 \{SF PROP-OR-QUES, E.ASPECT PERFECTIVE\},$

$h_3: \text{"_túk<sup>w</sup>_v_go.home_rel"}(ARG0: e_2, ARG1: x_4 \{PNG.PER 3RD\}),$

$h_3: \text{"_txw_x_ecs_rel"}(ARG0: e_2, ARG1: x_5 \{PNG.PER 3RD, PNG.GEND FEM\}),$

$h_6: \text{"_ti_q_unique_rel"}(ARG0: x_4, RSTR: h_7, BODY: h_8),$

$h_9: \text{"_čáčas_n_child_rel"}(ARG0: x_4),$

$h_{10}: \text{"_tsi_q_unique.F_rel"}(ARG0: x_5, RSTR: h_{11}, BODY: h_{12}),$

$h_{13}: \text{"_čáčas_n_child_rel"}(ARG0: x_5)$

$\{ h_7 =_q h_9, h_{11} =_q h_{13} \} \rangle$

1640 *ʔu-túk<sup>w</sup>-tx<sup>w</sup>-b čəx<sup>w</sup> ʔu ʔə ti luλ'*

$\langle h_1, e_2 \{SF QUES, E.ASPECT PERFECTIVE\},$

$h_3: \text{"_túk<sup>w</sup>_v_go.home_rel"}(ARG0: e_2, ARG1: x_4 \{PNG.PER 2ND, PNG.NUM SG\}),$

$h_3: \text{"_txw_x_ecs_rel"}(ARG0: e_2, ARG1: x_5 \{PNG.PER 3RD\}),$

$h_6: \text{"_ti_q_unique_rel"}(ARG0: x_5, RSTR: h_7, BODY: h_8),$

$h_9: \text{"_luλ'_n_older-person_rel"}(ARG0: x_5)$

$\{ h_7 =_q h_9 \} \rangle$

1650 *ʔu-təlawil-s-b čəd ʔə ti sqʷəbay?*

$\langle h_1, e_2 \{E.ASPECT PERFECTIVE\},$

$h_3: \text{"_təlawil\_v\_run\_rel"}(ARG0: e_2, ARG1: x_4 \{PNG.PER 3RD\}),$

$h_3: \text{"_c\_x\_altv\_rel"}(ARG0: e_2, ARG1: x_5 \{PNG.PER 1ST, PNG.NUM SG\}),$

$h_6: \text{"_ti\_q\_unique\_rel"}(ARG0: x_4, RSTR: h_7, BODY: h_8),$

$h_9: \text{"_sqʷəbay?_n\_dog\_rel"}(ARG0: x_4)$

$\{ h_7 =_q h_9 \} \rangle$

1660 *ʔu-gʷəč-t-b čət ʔə ti sqʷəbay?*

$\langle h_1, e_2 \{E.ASPECT PERFECTIVE\},$

$h_3: \text{"_gʷəč\_v\_look.for\_rel"}(ARG0: e_2, ARG1: x_4 \{PNG.PER 1ST, PNG.NUM PL\}),$

$h_3: \text{"_t\_x\_ics\_rel"}(ARG0: e_2, ARG1: x_5 \{PNG.PER 3RD\}),$

$h_6: \text{"_ti\_q\_unique\_rel"}(ARG0: x_5, RSTR: h_7, BODY: h_8),$

$h_9: \text{"_sqʷəbay?_n\_dog\_rel"}(ARG0: x_5)$

$\{ h_7 =_q h_9 \} \rangle$

1670 *ʔu-ʔəy-dxʷ-b čələp ʔu ʔə tsi čačas*

$\langle h_1, e_2 \{SF QUES, E.ASPECT PERFECTIVE\},$

$h_3: \text{"_ʔəy\_v\_found\_rel"}(ARG0: e_2, ARG1: x_4 \{PNG.PER 2ND, PNG.NUM PL\}),$

$h_3: \text{"_dxw\_x\_dc\_rel"}(ARG0: e_2, ARG1: x_5 \{PNG.PER 3RD, PNG.GEND FEM\}),$

$h_6: \text{"_tsi\_q\_unique.F\_rel"}(ARG0: x_5, RSTR: h_7, BODY: h_8),$

$h_9: \text{"_čačas\_n\_child\_rel"}(ARG0: x_5)$

$\{ h_7 =_q h_9 \} \rangle$

1680 *ʔu-ʔəł-c-b čəxʷ ʔu ʔə ti ʔaciʔtalbixʷ*

$\langle h_1, e_2 \{SF QUES, E.ASPECT PERFECTIVE\},$

$h_3: \text{"_ʔəł\_v\_come\_rel"}(ARG0: e_2, ARG1: x_4 \{PNG.PER 3RD\}),$

$h_3: \text{"_c\_x\_altv\_rel"}(ARG0: e_2, ARG1: x_5 \{PNG.PER 2ND, PNG.NUM SG\}),$

$h_6: \text{"_ti\_q\_unique\_rel"}(ARG0: x_4, RSTR: h_7, BODY: h_8),$

$h_9: \text{"_ʔaciʔtalbixʷ\_n\_person/indian\_rel"}(ARG0: x_4)$

$\{ h_7 =_q h_9 \} \rangle$

1690 *ʔu-təlawil-s-b ʔə ti sqʷəbayʔ ti luλ'*

⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE} ⟩,

$h_3$ : "\_təlawil\_v\_run\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),

$h_3$ : "\_c\_x\_altv\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD}),

$h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_7$ , BODY:  $h_8$ ),

$h_9$ : "\_sqʷəbayʔ\_n\_dog\_rel"(ARG0:  $x_4$ ),

$h_{10}$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_{11}$ , BODY:  $h_{12}$ ),

$h_{13}$ : "\_luλ'\_n\_older-person\_rel"(ARG0:  $x_5$ )

{  $h_7 =_q h_9, h_{11} =_q h_{13}$  } ⟩

1700 *ʔu-hədʔiʷ-c-b ʔə ti čačas tsi čačas*

⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE} ⟩,

$h_3$ : "\_hədʔiʷ\_v\_inside\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),

$h_3$ : "\_c\_x\_altv\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD, PNG.GEND FEM}),

$h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_7$ , BODY:  $h_8$ ),

$h_9$ : "\_čačas\_n\_child\_rel"(ARG0:  $x_4$ ),

$h_{10}$ : "\_tsi\_q\_unique.F\_rel"(ARG0:  $x_5$ , RSTR:  $h_{11}$ , BODY:  $h_{12}$ ),

$h_{13}$ : "\_čačas\_n\_child\_rel"(ARG0:  $x_5$ )

{  $h_7 =_q h_9, h_{11} =_q h_{13}$  } ⟩

1710 *ʔu-gʷəč-t-b ʔə tsi luλ' tsi čačas*

⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE} ⟩,

$h_3$ : "\_gʷəč\_v\_look.for\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD, PNG.GEND FEM}),

$h_3$ : "\_t\_x\_ics\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD, PNG.GEND FEM}),

$h_6$ : "\_tsi\_q\_unique.F\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),

$h_9$ : "\_luλ'\_n\_older-person\_rel"(ARG0:  $x_5$ ),

$h_{10}$ : "\_tsi\_q\_unique.F\_rel"(ARG0:  $x_4$ , RSTR:  $h_{11}$ , BODY:  $h_{12}$ ),

$h_{13}$ : "\_čačas\_n\_child\_rel"(ARG0:  $x_4$ )

{  $h_7 =_q h_9, h_{11} =_q h_{13}$  } ⟩

1720 *ʔu-ʔəʔʰ-tx<sup>w</sup>-b ʔə ti ʔaciʔtalbix<sup>w</sup> ti luʔ* $\langle h_1, e_2 \{E.ASPECT PERFECTIVE\},$  $h_3: \text{"_ʔəʔʰ_v_come\_rel"}(ARG0: e_2, ARG1: x_4 \{PNG.PER 3RD\}),$  $h_3: \text{"_txw\_x\_ecs\_rel"}(ARG0: e_2, ARG1: x_5 \{PNG.PER 3RD\}),$  $h_6: \text{"_ti\_q\_unique\_rel"}(ARG0: x_5, RSTR: h_7, BODY: h_8),$  $h_9: \text{"_ʔaciʔtalbix<sup>w</sup>\_n\_person/indian\_rel"}(ARG0: x_5),$  $h_{10}: \text{"_ti\_q\_unique\_rel"}(ARG0: x_4, RSTR: h_{11}, BODY: h_{12}),$  $h_{13}: \text{"_luʔ\_n\_older-person\_rel"}(ARG0: x_4)$  $\{ h_7 =_q h_9, h_{11} =_q h_{13} \} \rangle$ 1730 *ʔu-tʰuk<sup>w</sup>-tx<sup>w</sup>-b čəx<sup>w</sup> ʔu ʔə tsi luʔ* $\langle h_1, e_2 \{SF QUES, E.ASPECT PERFECTIVE\},$  $h_3: \text{"_tʰuk<sup>w</sup>\_v\_go.home\_rel"}(ARG0: e_2, ARG1: x_4 \{PNG.PER 2ND, PNG.NUM SG\}),$  $h_3: \text{"_txw\_x\_ecs\_rel"}(ARG0: e_2, ARG1: x_5 \{PNG.PER 3RD, PNG.GEND FEM\}),$  $h_6: \text{"_tsi\_q\_unique.F\_rel"}(ARG0: x_5, RSTR: h_7, BODY: h_8),$  $h_9: \text{"_luʔ\_n\_older-person\_rel"}(ARG0: x_5)$  $\{ h_7 =_q h_9 \} \rangle$ 1740 *ʔu-ʔəy<sup>ʷ</sup>-dx<sup>w</sup>-b čəd ʔə ti čačas* $\langle h_1, e_2 \{E.ASPECT PERFECTIVE\},$  $h_3: \text{"_ʔəy<sup>ʷ</sup>\_v\_found\_rel"}(ARG0: e_2, ARG1: x_4 \{PNG.PER 1ST, PNG.NUM SG\}),$  $h_3: \text{"_dxw\_x\_dc\_rel"}(ARG0: e_2, ARG1: x_5 \{PNG.PER 3RD\}),$  $h_6: \text{"_ti\_q\_unique\_rel"}(ARG0: x_5, RSTR: h_7, BODY: h_8),$  $h_9: \text{"_čačas\_n\_child\_rel"}(ARG0: x_5)$  $\{ h_7 =_q h_9 \} \rangle$ 1750 *ʔu-ʔčil-s-b čəʔ ʔə ti ʔaciʔtalbix<sup>w</sup>* $\langle h_1, e_2 \{E.ASPECT PERFECTIVE\},$  $h_3: \text{"_ʔčil\_v\_arrive\_rel"}(ARG0: e_2, ARG1: x_4 \{PNG.PER 3RD\}),$  $h_3: \text{"_c\_x\_altv\_rel"}(ARG0: e_2, ARG1: x_5 \{PNG.PER 1ST, PNG.NUM PL\}),$  $h_6: \text{"_ti\_q\_unique\_rel"}(ARG0: x_4, RSTR: h_7, BODY: h_8),$  $h_9: \text{"_ʔaciʔtalbix<sup>w</sup>\_n\_person/indian\_rel"}(ARG0: x_4)$  $\{ h_7 =_q h_9 \} \rangle$

1760 *ʔu-čubə-c-b čələp ʔu ʔə ti luʔ*

⟨  $h_1, e_2$  {SF QUES, E.ASPECT PERFECTIVE},  
 $h_3$ : "\_čubə\_v\_go.inland\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),  
 $h_3$ : "\_c\_x\_altv\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 2ND, PNG.NUM PL}),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_luʔ\_n\_older-person\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9$  } ⟩

1770 *ʔu-bəč-t-b čəx<sup>w</sup> ʔu ʔə tsi luʔ*

⟨  $h_1, e_2$  {SF QUES, E.ASPECT PERFECTIVE},  
 $h_3$ : "\_bəč\_v\_fallen\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 2ND, PNG.NUM SG}),  
 $h_3$ : "\_t\_x\_ics\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD, PNG.GEND FEM}),  
 $h_6$ : "\_tsi\_q\_unique.F\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_luʔ\_n\_older-person\_rel"(ARG0:  $x_5$ )  
 {  $h_7 =_q h_9$  } ⟩

1780 *ʔu-g<sup>w</sup>əč-t tsi čačas*

⟨  $h_1, e_2$  {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 $h_3$ : "\_g<sup>w</sup>əč\_v\_look.for\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD, PNG.GEND FEM}),  
 $h_3$ : "\_t\_x\_ics\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$ ),  
 $h_6$ : "\_tsi\_q\_unique.F\_rel"(ARG0:  $x_4$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_čačas\_n\_child\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9$  } ⟩

1790 *ʔu-g<sup>w</sup>əč-b tsi čačas*

⟨  $h_1, e_2$  {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 $h_3$ : "\_g<sup>w</sup>əč\_v\_look.for\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ ),  
 $h_3$ : "\_b\_x\_csmd\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD, PNG.GEND FEM}),  
 $h_6$ : "\_tsi\_q\_unique.F\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_čačas\_n\_child\_rel"(ARG0:  $x_5$ )  
 {  $h_7 =_q h_9$  } ⟩

1800 *ʔu-ǰʷəl-b tsi čačas*

⟨  $h_1, e_2$  {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 $h_3$ : "\_ǰʷəl\_v\_cooked/ripened\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ ),  
 $h_3$ : "\_b\_x\_csmd\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD, PNG.GEND FEM}),  
 $h_6$ : "\_tsi\_q\_unique.F\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_čačas\_n\_child\_rel"(ARG0:  $x_5$ )  
 {  $h_7 =_q h_9$  } ⟩

1810 *ʔu-ǰʷəl-t ti sʔuladxʷ*

⟨  $h_1, e_2$  {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 $h_3$ : "\_ǰʷəl\_v\_cooked/ripened\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),  
 $h_3$ : "\_t\_x\_ics\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$ ),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_sʔuladxʷ\_n\_salmon\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9$  } ⟩

1820 *ʔu-hədʔiʷ-t ti sqʷəbayʔ*

⟨  $h_1, e_2$  {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 $h_3$ : "\_hədʔiʷ\_v\_inside\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),  
 $h_3$ : "\_t\_x\_ics\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$ ),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_sqʷəbayʔ\_n\_dog\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9$  } ⟩

1830 *ʔu-hədʔiʷ-b ti sqʷəbayʔ*

⟨  $h_1, e_2$  {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 $h_3$ : "\_hədʔiʷ\_v\_inside\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ ),  
 $h_3$ : "\_b\_x\_csmd\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD}),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_sqʷəbayʔ\_n\_dog\_rel"(ARG0:  $x_5$ )  
 {  $h_7 =_q h_9$  } ⟩

1840 *ʔu-gʷəč̣-t čəd tsi čačas*

$\langle h_1, e_2 \{E.ASPECT PERFECTIVE\},$   
 $h_3: \text{"_gʷəč̣_v\_look.for\_rel"}(ARG0: e_2, ARG1: x_4 \{PNG.PER 3RD, PNG.GEND FEM\}),$   
 $h_3: \text{"_t\_x\_ics\_rel"}(ARG0: e_2, ARG1: x_5 \{PNG.PER 1ST, PNG.NUM SG\}),$   
 $h_6: \text{"_tsi\_q\_unique.F\_rel"}(ARG0: x_4, RSTR: h_7, BODY: h_8),$   
 $h_9: \text{"_čačas\_n\_child\_rel"}(ARG0: x_4)$   
 $\{ h_7 =_q h_9 \} \rangle$

1850 *ʔu-gʷəč̣-b čəd*

$\langle h_1, e_2 \{E.ASPECT PERFECTIVE\},$   
 $h_3: \text{"_gʷəč̣_v\_look.for\_rel"}(ARG0: e_2, ARG1: x_4),$   
 $h_3: \text{"_b\_x\_csmd\_rel"}(ARG0: e_2, ARG1: x_5 \{PNG.PER 1ST, PNG.NUM SG\})$   
 $\{ \} \rangle$

1860 *ʔu-ǰʷəl-b čəd*

$\langle h_1, e_2 \{E.ASPECT PERFECTIVE\},$   
 $h_3: \text{"_ǰʷəl\_v\_cooked/ripened\_rel"}(ARG0: e_2, ARG1: x_4),$   
 $h_3: \text{"_b\_x\_csmd\_rel"}(ARG0: e_2, ARG1: x_5 \{PNG.PER 1ST, PNG.NUM SG\})$   
 $\{ \} \rangle$

1870 *ʔu-ǰʷəl-t čəl ti sʔuladxʷ*

$\langle h_1, e_2 \{E.ASPECT PERFECTIVE\},$   
 $h_3: \text{"_ǰʷəl\_v\_cooked/ripened\_rel"}(ARG0: e_2, ARG1: x_4 \{PNG.PER 3RD\}),$   
 $h_3: \text{"_t\_x\_ics\_rel"}(ARG0: e_2, ARG1: x_5 \{PNG.PER 1ST, PNG.NUM PL\}),$   
 $h_6: \text{"_ti\_q\_unique\_rel"}(ARG0: x_4, RSTR: h_7, BODY: h_8),$   
 $h_9: \text{"_sʔuladxʷ\_n\_salmon\_rel"}(ARG0: x_4)$   
 $\{ h_7 =_q h_9 \} \rangle$

1880 *ʔu-hədʔiʷ-t čəxʷ ʔu ti sqʷəbayʔ*

$\langle h_1, e_2 \{SF QUES, E.ASPECT PERFECTIVE\},$   
 $h_3: \text{"_hədʔiʷ\_v\_inside\_rel"}(ARG0: e_2, ARG1: x_4 \{PNG.PER 3RD\}),$   
 $h_3: \text{"_t\_x\_ics\_rel"}(ARG0: e_2, ARG1: x_5 \{PNG.PER 2ND, PNG.NUM SG\}),$   
 $h_6: \text{"_ti\_q\_unique\_rel"}(ARG0: x_4, RSTR: h_7, BODY: h_8),$   
 $h_9: \text{"_sqʷəbayʔ\_n\_dog\_rel"}(ARG0: x_4)$   
 $\{ h_7 =_q h_9 \} \rangle$

1890 *ʔu-hədʔiw-b čələp ʔu*

$\langle h_1, e_2 \{ \text{SF QUES, E.ASPECT PERFECTIVE} \},$   
 $h_3: \text{"_hədʔiw\_v\_inside\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_4),$   
 $h_3: \text{"_b\_x\_csmd\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_5 \{ \text{PNG.PER 2ND, PNG.NUM PL} \})$   
 $\{ \}$

1900 *ʔu-gʷəč-t čəd ti sqʷəbay?*

$\langle h_1, e_2 \{ \text{E.ASPECT PERFECTIVE} \},$   
 $h_3: \text{"_gʷəč\_v\_look.for\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_4 \{ \text{PNG.PER 3RD} \}),$   
 $h_3: \text{"_t\_x\_ics\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_5 \{ \text{PNG.PER 1ST, PNG.NUM SG} \}),$   
 $h_6: \text{"_ti\_q\_unique\_rel"}(\text{ARG0: } x_4, \text{RSTR: } h_7, \text{BODY: } h_8),$   
 $h_9: \text{"_sqʷəbay?_n\_dog\_rel"}(\text{ARG0: } x_4)$   
 $\{ h_7 =_q h_9 \}$

1910 *ʔu-gʷəč-t čət ti sqʷəbay?*

$\langle h_1, e_2 \{ \text{E.ASPECT PERFECTIVE} \},$   
 $h_3: \text{"_gʷəč\_v\_look.for\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_4 \{ \text{PNG.PER 3RD} \}),$   
 $h_3: \text{"_t\_x\_ics\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_5 \{ \text{PNG.PER 1ST, PNG.NUM PL} \}),$   
 $h_6: \text{"_ti\_q\_unique\_rel"}(\text{ARG0: } x_4, \text{RSTR: } h_7, \text{BODY: } h_8),$   
 $h_9: \text{"_sqʷəbay?_n\_dog\_rel"}(\text{ARG0: } x_4)$   
 $\{ h_7 =_q h_9 \}$

1920 *ʔu-gʷəč-t čəxʷ ti sqʷəbay?*

$\langle h_1, e_2 \{ \text{E.ASPECT PERFECTIVE} \},$   
 $h_3: \text{"_gʷəč\_v\_look.for\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_4 \{ \text{PNG.PER 3RD} \}),$   
 $h_3: \text{"_t\_x\_ics\_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_5 \{ \text{PNG.PER 2ND, PNG.NUM SG} \}),$   
 $h_6: \text{"_ti\_q\_unique\_rel"}(\text{ARG0: } x_4, \text{RSTR: } h_7, \text{BODY: } h_8),$   
 $h_9: \text{"_sqʷəbay?_n\_dog\_rel"}(\text{ARG0: } x_4)$   
 $\{ h_7 =_q h_9 \}$

1930 *ʔu-gʷəč̣-t čələp ti sqʷəbay?*

⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},  
 $h_3$ : "\_gʷəč̣\_v\_look.for\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),  
 $h_3$ : "\_t\_x\_ics\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 2ND, PNG.NUM PL}),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_sqʷəbay?\_n\_dog\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9$  } ⟩

1940 *ʔu-gʷəč̣-t ti sqʷəbay?*

⟨  $h_1, e_2$  {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 $h_3$ : "\_gʷəč̣\_v\_look.for\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),  
 $h_3$ : "\_t\_x\_ics\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$ ),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_sqʷəbay?\_n\_dog\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9$  } ⟩

1950 *ʔu-gʷəč̣-t čəd*

⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},  
 $h_3$ : "\_gʷəč̣\_v\_look.for\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ ),  
 $h_3$ : "\_t\_x\_ics\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 1ST, PNG.NUM SG})  
 { } ⟩

1960 *ʔu-gʷəč̣-b čəd*

⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},  
 $h_3$ : "\_gʷəč̣\_v\_look.for\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ ),  
 $h_3$ : "\_b\_x\_csmd\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 1ST, PNG.NUM SG})  
 { } ⟩

1970 *ʔu-ʔuχ̣ʷ čəd*

⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},  
 $h_3$ : "\_ʔuχ̣ʷ\_v\_go\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 1ST, PNG.NUM SG})  
 { } ⟩

- 1980 *ʔu-ʔuχ<sup>w</sup>-tx<sup>w</sup> čəd ti čáčas*  
 ⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},  
 $h_3$ : "\_ʔuχ<sup>w</sup>\_v\_go\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ {PNG.PER 3RD}),  
 $h_3$ : "\_txw\_x\_ecs\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$ {PNG.PER 1ST, PNG.NUM SG}),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_čáčas\_n\_child\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9$  } ⟩
- 1990 *ʔu-ʔəł<sup>w</sup>-tx<sup>w</sup> čəd ti čáčas*  
 ⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},  
 $h_3$ : "\_ʔəł<sup>w</sup>\_v\_come\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ {PNG.PER 3RD}),  
 $h_3$ : "\_txw\_x\_ecs\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$ {PNG.PER 1ST, PNG.NUM SG}),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_čáčas\_n\_child\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9$  } ⟩
- 2000 *ʔu-ʔúk<sup>w</sup> čət*  
 ⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},  
 $h_3$ : "\_ʔúk<sup>w</sup>\_v\_go.home\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ {PNG.PER 1ST, PNG.NUM PL})  
 { } ⟩
- 2010 *ʔu-ʔúk<sup>w</sup>-tx<sup>w</sup> čət ti čáčas*  
 ⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},  
 $h_3$ : "\_ʔúk<sup>w</sup>\_v\_go.home\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ {PNG.PER 3RD}),  
 $h_3$ : "\_txw\_x\_ecs\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$ {PNG.PER 1ST, PNG.NUM PL}),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_čáčas\_n\_child\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9$  } ⟩
- 2020 *ʔu-ʔəł<sup>w</sup> čət*  
 ⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},  
 $h_3$ : "\_ʔəł<sup>w</sup>\_v\_come\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ {PNG.PER 1ST, PNG.NUM PL})  
 { } ⟩
- 2030 *ʔu-ʔəł<sup>w</sup> čəx<sup>w</sup>*  
 ⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},  
 $h_3$ : "\_ʔəł<sup>w</sup>\_v\_come\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ {PNG.PER 2ND, PNG.NUM SG})  
 { } ⟩

- 2040 *ʔu-ʔuχ<sup>w</sup> čəx<sup>w</sup> ʔu*  
 ⟨  $h_1, e_2$ {SF QUES, E.ASPECT PERFECTIVE},  
 |  $h_3$ :"\_ʔuχ<sup>w</sup>\_v\_go\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ {PNG.PER 2ND, PNG.NUM SG}) |  
 { }⟩
- 2050 *ʔu-ʔuk<sup>w</sup>-tx<sup>w</sup> čəx<sup>w</sup> ʔu ti čačas*  
 ⟨  $h_1, e_2$ {SF QUES, E.ASPECT PERFECTIVE},  
 |  $h_3$ :"\_ʔuk<sup>w</sup>\_v\_go.home\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ {PNG.PER 3RD}),  
 |  $h_3$ :"\_txw\_x\_ecs\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$ {PNG.PER 2ND, PNG.NUM SG}),  
 |  $h_6$ :"\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 |  $h_9$ :"\_čačas\_n\_child\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9$  }⟩
- 2060 *ʔu-ʔuχ<sup>w</sup>-tx<sup>w</sup> čələp ʔu ti čačas*  
 ⟨  $h_1, e_2$ {SF QUES, E.ASPECT PERFECTIVE},  
 |  $h_3$ :"\_ʔuχ<sup>w</sup>\_v\_go\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ {PNG.PER 3RD}),  
 |  $h_3$ :"\_txw\_x\_ecs\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$ {PNG.PER 2ND, PNG.NUM PL}),  
 |  $h_6$ :"\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 |  $h_9$ :"\_čačas\_n\_child\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9$  }⟩
- 2070 *ʔu-ʔəλ' čələp*  
 ⟨  $h_1, e_2$ {E.ASPECT PERFECTIVE},  
 |  $h_3$ :"\_ʔəλ'\_v\_come\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ {PNG.PER 2ND, PNG.NUM PL}) |  
 { }⟩
- 2080 *ʔu-ʔuχ<sup>w</sup> čələp ʔu*  
 ⟨  $h_1, e_2$ {SF QUES, E.ASPECT PERFECTIVE},  
 |  $h_3$ :"\_ʔuχ<sup>w</sup>\_v\_go\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ {PNG.PER 2ND, PNG.NUM PL}) |  
 { }⟩
- 2090 *ʔu-ʔəλ' ʔu ti čačas*  
 ⟨  $h_1, e_2$ {SF QUES, E.ASPECT PERFECTIVE},  
 |  $h_3$ :"\_ʔəλ'\_v\_come\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ {PNG.PER 3RD}),  
 |  $h_5$ :"\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_6$ , BODY:  $h_7$ ),  
 |  $h_8$ :"\_čačas\_n\_child\_rel"(ARG0:  $x_4$ )  
 {  $h_6 =_q h_8$  }⟩

2100 *ʔu-ʔəʔʰ-txʷ ʔu ti čačas*

$\langle h_1, e_2 \{ \text{SF QUES, E.ASPECT PERFECTIVE} \},$   
 $h_3: \text{"_ʔəʔʰ_v_come_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_4 \{ \text{PNG.PER 3RD} \}),$   
 $h_3: \text{"_txw_x_ecs_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_5),$   
 $h_6: \text{"_ti_q_unique_rel"}(\text{ARG0: } x_4, \text{RSTR: } h_7, \text{BODY: } h_8),$   
 $h_9: \text{"_čačas_n_child_rel"}(\text{ARG0: } x_4)$   
 $\{ h_7 =_q h_9 \} \rangle$

2110 *ʔu-ʔuχʷ tsi čačas*

$\langle h_1, e_2 \{ \text{SF PROP-OR-QUES, E.ASPECT PERFECTIVE} \},$   
 $h_3: \text{"_ʔuχʷ_v_go_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_4 \{ \text{PNG.PER 3RD, PNG.GEND FEM} \}),$   
 $h_5: \text{"_tsi_q_unique.F_rel"}(\text{ARG0: } x_4, \text{RSTR: } h_6, \text{BODY: } h_7),$   
 $h_8: \text{"_čačas_n_child_rel"}(\text{ARG0: } x_4)$   
 $\{ h_6 =_q h_8 \} \rangle$

2120 *ʔu-ʔusil ti čačas*

$\langle h_1, e_2 \{ \text{SF PROP-OR-QUES, E.ASPECT PERFECTIVE} \},$   
 $h_3: \text{"_ʔusil_v_dive_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_4 \{ \text{PNG.PER 3RD} \}),$   
 $h_5: \text{"_ti_q_unique_rel"}(\text{ARG0: } x_4, \text{RSTR: } h_6, \text{BODY: } h_7),$   
 $h_8: \text{"_čačas_n_child_rel"}(\text{ARG0: } x_4)$   
 $\{ h_6 =_q h_8 \} \rangle$

2130 *ʔu-ʔusil čəd*

$\langle h_1, e_2 \{ \text{E.ASPECT PERFECTIVE} \},$   
 $h_3: \text{"_ʔusil_v_dive_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_4 \{ \text{PNG.PER 1ST, PNG.NUM SG} \})$   
 $\{ \} \rangle$

2140 *ʔu-saxʷəb ti čačas*

$\langle h_1, e_2 \{ \text{SF PROP-OR-QUES, E.ASPECT PERFECTIVE} \},$   
 $h_3: \text{"_saxʷəb_v_jump_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_4 \{ \text{PNG.PER 3RD} \}),$   
 $h_5: \text{"_ti_q_unique_rel"}(\text{ARG0: } x_4, \text{RSTR: } h_6, \text{BODY: } h_7),$   
 $h_8: \text{"_čačas_n_child_rel"}(\text{ARG0: } x_4)$   
 $\{ h_6 =_q h_8 \} \rangle$

2150 *ʔu-saxʷəb čəd*

$\langle h_1, e_2 \{ \text{E.ASPECT PERFECTIVE} \},$   
 $h_3: \text{"_saxʷəb_v_jump_rel"}(\text{ARG0: } e_2, \text{ARG1: } x_4 \{ \text{PNG.PER 1ST, PNG.NUM SG} \})$   
 $\{ \} \rangle$

2160 *ʔu-gʷəč̣-b ti luʔʔə ti sqəlalitut*

⟨  $h_1, e_2$  {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 $h_3$ : "\_gʷəč̣\_v\_look.for\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD}),  
 $h_3$ : "\_b\_x\_csmd\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD}),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_luʔʔ\_n\_older-person\_rel"(ARG0:  $x_5$ ),  
 $h_{10}$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_{11}$ , BODY:  $h_{12}$ ),  
 $h_{13}$ : "\_sqəlalitut\_n\_spirit.power\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9, h_{11} =_q h_{13}$  } ⟩

2170 *ʔu-ʔəʔəd ti ʔacitʔalbixʷ ʔə ti sʔuladxʷ*

⟨  $h_1, e_2$  {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 $h_3$ : "\_ʔəʔəd\_v\_eat\_something\_rel"(ARG0:  $e_2$ ,  
 ARG1:  $x_5$  {PNG.PER 3RD},  
 ARG2:  $x_4$  {PNG.PER 3RD}),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_ʔacitʔalbixʷ\_n\_person/indian\_rel"(ARG0:  $x_5$ ),  
 $h_{10}$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_{11}$ , BODY:  $h_{12}$ ),  
 $h_{13}$ : "\_sʔuladxʷ\_n\_salmon\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9, h_{11} =_q h_{13}$  } ⟩

2180 *ʔu-ʔəʔəd čəd ʔə ti sʔuladxʷ*

⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},  
 $h_3$ : "\_ʔəʔəd\_v\_eat\_something\_rel"(ARG0:  $e_2$ ,  
 ARG1:  $x_5$  {PNG.PER 1ST, PNG.NUM SG},  
 ARG2:  $x_4$  {PNG.PER 3RD}),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_sʔuladxʷ\_n\_salmon\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9$  } ⟩

- 2190 *ʔu-ʔuχ<sup>w</sup>-tx<sup>w</sup> tsi čačas*  
 ⟨  $h_1, e_2$  {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 $h_3$ : "\_ʔuχ<sup>w</sup>\_v\_go\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ {PNG.PER 3RD, PNG.GEND FEM}),  
 $h_3$ : "\_txw\_x\_ecs\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$ ),  
 $h_6$ : "\_tsi\_q\_unique.F\_rel"(ARG0:  $x_4$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_čačas\_n\_child\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9$  } ⟩
- 2200 *ʔu-ʔuχ<sup>w</sup>-tx<sup>w</sup> čad*  
 ⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},  
 $h_3$ : "\_ʔuχ<sup>w</sup>\_v\_go\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ ),  
 $h_3$ : "\_txw\_x\_ecs\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$ {PNG.PER 1ST, PNG.NUM SG})  
 { } ⟩
- 2210 *ʔu-ʔuχ<sup>w</sup>-c tsi čačas*  
 ⟨  $h_1, e_2$  {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 $h_3$ : "\_ʔuχ<sup>w</sup>\_v\_go\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ ),  
 $h_3$ : "\_c\_x\_altv\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$ {PNG.PER 3RD, PNG.GEND FEM}),  
 $h_6$ : "\_tsi\_q\_unique.F\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_čačas\_n\_child\_rel"(ARG0:  $x_5$ )  
 {  $h_7 =_q h_9$  } ⟩
- 2220 *ʔu-ʔuχ<sup>w</sup>-c čad*  
 ⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},  
 $h_3$ : "\_ʔuχ<sup>w</sup>\_v\_go\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ {PNG.PER 1ST, PNG.NUM SG}),  
 $h_3$ : "\_c\_x\_altv\_rel"(ARG0:  $e_2$ , ARG1:  $i_5$ )  
 { } ⟩
- 2230 *ʔu-ʔusil-s ti čačas*  
 ⟨  $h_1, e_2$  {SF PROP-OR-QUES, E.ASPECT PERFECTIVE},  
 $h_3$ : "\_ʔusil\_v\_dive\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ ),  
 $h_3$ : "\_c\_x\_altv\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$ {PNG.PER 3RD}),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_čačas\_n\_child\_rel"(ARG0:  $x_5$ )  
 {  $h_7 =_q h_9$  } ⟩

2240 *ʔu-ʔusil-s čəd*

$\langle h_1, e_2 \{E.ASPECT PERFECTIVE\},$   
 $h_3: \_ʔusil\_v\_dive\_rel($ ARG0:  $e_2$ , ARG1:  $x_4 \{PNG.PER 1ST, PNG.NUM SG\}$ ),  
 $h_3: \_c\_x\_altv\_rel($ ARG0:  $e_2$ , ARG1:  $i_5$ )  
 $\{ \} \rangle$

2250 *ʔu-ʔəy̌-dx<sup>w</sup> ti čáčas*

$\langle h_1, e_2 \{SF PROP-OR-QUES, E.ASPECT PERFECTIVE\},$   
 $h_3: \_ʔəy̌\_v\_found\_rel($ ARG0:  $e_2$ , ARG1:  $x_4 \{PNG.PER 3RD\}$ ),  
 $h_3: \_dxw\_x\_dc\_rel($ ARG0:  $e_2$ , ARG1:  $x_5$ ),  
 $h_6: \_ti\_q\_unique\_rel($ ARG0:  $x_4$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9: \_čáčas\_n\_child\_rel($ ARG0:  $x_4$ )  
 $\{ h_7 =_q h_9 \} \rangle$

2260 *ʔu-ʔəy̌-dx<sup>w</sup> čəd*

$\langle h_1, e_2 \{E.ASPECT PERFECTIVE\},$   
 $h_3: \_ʔəy̌\_v\_found\_rel($ ARG0:  $e_2$ , ARG1:  $x_4$ ),  
 $h_3: \_dxw\_x\_dc\_rel($ ARG0:  $e_2$ , ARG1:  $x_5 \{PNG.PER 1ST, PNG.NUM SG\}$ )  
 $\{ \} \rangle$

2270 *ʔu-ʔuχ<sup>w</sup>-tx<sup>w</sup>-b ʔə ti čáčas tsi čáčas*

$\langle h_1, e_2 \{E.ASPECT PERFECTIVE\},$   
 $h_3: \_ʔuχ<sup>w</sup>\_v\_go\_rel($ ARG0:  $e_2$ , ARG1:  $x_4 \{PNG.PER 3RD, PNG.GEND FEM\}$ ),  
 $h_3: \_txw\_x\_ecs\_rel($ ARG0:  $e_2$ , ARG1:  $x_5 \{PNG.PER 3RD\}$ ),  
 $h_6: \_ti\_q\_unique\_rel($ ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9: \_čáčas\_n\_child\_rel($ ARG0:  $x_5$ ),  
 $h_{10}: \_tsi\_q\_unique.F\_rel($ ARG0:  $x_4$ , RSTR:  $h_{11}$ , BODY:  $h_{12}$ ),  
 $h_{13}: \_čáčas\_n\_child\_rel($ ARG0:  $x_4$ )  
 $\{ h_7 =_q h_9, h_{11} =_q h_{13} \} \rangle$

2280 *ʔu-ʔuχ<sup>w</sup>-tx<sup>w</sup>-b čəd ʔə ti čáčas*

⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},  
 $h_3$ : "\_ʔuχ<sup>w</sup>\_v\_go\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ {PNG.PER 1ST, PNG.NUM SG}),  
 $h_3$ : "\_txw\_x\_ecs\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$ {PNG.PER 3RD}),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_čáčas\_n\_child\_rel"(ARG0:  $x_5$ )  
 {  $h_7 =_q h_9$  } ⟩

2290 *ʔu-ʔuχ<sup>w</sup>-c-b ʔə ti čáčas tsi čáčas*

⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},  
 $h_3$ : "\_ʔuχ<sup>w</sup>\_v\_go\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ {PNG.PER 3RD}),  
 $h_3$ : "\_c\_x\_altv\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$ {PNG.PER 3RD, PNG.GEND FEM}),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_čáčas\_n\_child\_rel"(ARG0:  $x_4$ ),  
 $h_{10}$ : "\_tsi\_q\_unique.F\_rel"(ARG0:  $x_5$ , RSTR:  $h_{11}$ , BODY:  $h_{12}$ ),  
 $h_{13}$ : "\_čáčas\_n\_child\_rel"(ARG0:  $x_5$ )  
 {  $h_7 =_q h_9, h_{11} =_q h_{13}$  } ⟩

2300 *ʔu-ʔuχ<sup>w</sup>-c-b čəd ʔə ti čáčas*

⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},  
 $h_3$ : "\_ʔuχ<sup>w</sup>\_v\_go\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ {PNG.PER 3RD}),  
 $h_3$ : "\_c\_x\_altv\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$ {PNG.PER 1ST, PNG.NUM SG}),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_čáčas\_n\_child\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9$  } ⟩

2310 *ʔu-ʔusil-s-b čəd ʔə ti čičičiχ*

⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},  
 $h_3$ : "\_ʔusil\_v\_dive\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$ {PNG.PER 3RD}),  
 $h_3$ : "\_c\_x\_altv\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$ {PNG.PER 1ST, PNG.NUM SG}),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_4$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_čičičiχ<sup>w</sup>\_n\_osprey\_rel"(ARG0:  $x_4$ )  
 {  $h_7 =_q h_9$  } ⟩

2320 *ʔu-gʷəč-t-b čəd ʔə ti čačas*

⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},  
 $h_3$ : "\_gʷəč\_v\_look.for\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 1ST, PNG.NUM SG}),  
 $h_3$ : "\_t\_x\_ics\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD}),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_čačas\_n\_child\_rel"(ARG0:  $x_5$ )  
 $\{ h_7 =_q h_9 \}$  ⟩

2330 *ʔu-ʔəy-dxʷ-b ʔə ti čačas tsi čačas*

⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},  
 $h_3$ : "\_ʔəy\_v\_found\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 3RD, PNG.GEND FEM}),  
 $h_3$ : "\_dxw\_x\_dc\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD}),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_čačas\_n\_child\_rel"(ARG0:  $x_5$ ),  
 $h_{10}$ : "\_tsi\_q\_unique.F\_rel"(ARG0:  $x_4$ , RSTR:  $h_{11}$ , BODY:  $h_{12}$ ),  
 $h_{13}$ : "\_čačas\_n\_child\_rel"(ARG0:  $x_4$ )  
 $\{ h_7 =_q h_9, h_{11} =_q h_{13} \}$  ⟩

2340 *ʔu-ʔəy-dxʷ-b čəd ʔə ti čačas*

⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},  
 $h_3$ : "\_ʔəy\_v\_found\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 1ST, PNG.NUM SG}),  
 $h_3$ : "\_dxw\_x\_dc\_rel"(ARG0:  $e_2$ , ARG1:  $x_5$  {PNG.PER 3RD}),  
 $h_6$ : "\_ti\_q\_unique\_rel"(ARG0:  $x_5$ , RSTR:  $h_7$ , BODY:  $h_8$ ),  
 $h_9$ : "\_čačas\_n\_child\_rel"(ARG0:  $x_5$ )  
 $\{ h_7 =_q h_9 \}$  ⟩

2350 *ʔu-xʷitil čəd*

⟨  $h_1, e_2$  {E.ASPECT PERFECTIVE},  
 $h_3$ : "\_xʷitil\_v\_fall.off\_rel"(ARG0:  $e_2$ , ARG1:  $x_4$  {PNG.PER 1ST, PNG.NUM SG})  
 $\{ \}$  ⟩