

The title of this volume is shorter
than its contributions are
allowed to be:

Papers in honour of Hotze Rullmann

Edited by M. Ryan Bochnak,
Eva Csipak, Lisa Matthewson,
Marcin Morzycki, and
Daniel K. E. Reisinger

UBCOPL
2024

Copyright 2024, UBCOPL.

All rights to individual papers remain with the authors.

The title of this volume is shorter than its contributions are allowed to be: Papers in honour of Hotze Rullmann.

Edited by M. Ryan Bochnak, Eva Csipak, Lisa Matthewson, Marcin Morzycki, and Daniel K. E. Reisinger

Published by University of British Columbia Occasional Papers in Linguistics. First printing in 2024.

ISBN: 978-0-88865-492-2 UBCOPL Vol. 9

University of British Columbia
Occasional Papers in Linguistics
2613 West Mall
Vancouver, BC V6T 1Z4
Canada

For more information, please see lingpapers.sites.olt.ubc.ca.

Table of Contents

Preface <i>M. Ryan Bochnak, Eva Csipak, Lisa Matthewson, Marcin Morzycki, and Daniel K. E. Reisinger</i>	x
Degree constructions in Gitksan <i>Yurika Aonuki</i>	1
Incorporating sign language phonetics and phonology exercises into the Linguistics classroom <i>Yurika Aonuki and Kathleen Currie Hall</i>	19
Reversing the gaze: decolonizing the syllabus <i>Solveiga Armoskaite</i>	39
Celebrating and quantifying the linguistic diversity of the UBC student community <i>Molly Babel, Line Lloy, Zara Khalaji Pirbaluti, Rachel Soo, and Lexia Suite</i>	53
Mood in Wá·šiw narratives: A first look <i>M. Ryan Bochnak</i>	65
A second-last position clitic in Sm'algyax: A puzzle <i>Colin Brown and Henry Davis</i>	79
Evidence for two types of future semantics by negation <i>Sihwei Chen</i>	95
HOW many topics has Hotze worked on? Echo questions, mumbling and incredulity <i>Eva Csipak</i>	109
A second-last position clitic in Sm'algyax: A solution <i>Henry Davis and Colin Brown</i>	119
A pilot study: Affect and grammatical anomaly in discourse <i>Veena Dwivedi</i>	131

Enough!		
	<i>Kai von Fintel</i>	145
A sketch of content question formation in Eauripik Woleaian		
	<i>Laura Griffin</i>	157
<i>If I had half a brain</i> : Polarity sensitive idioms in conditional clauses		
	<i>Jack Hoeksema</i>	171
Variable strength in necessity modality: A case of variation between Afrikaans, Dutch, English, and German		
	<i>Vera Hohaus, Siena Weingartz, and Marijn Boomars</i>	183
On the absence of evidence for nominal tense: The ʔayʔajuθəm past marker -ol		
	<i>Marianne Huijsmans</i>	197
The passive in Heritage Icelandic		
	<i>Kristín M. Jóhannsdóttir</i>	217
Where to interpret what		
	<i>Kyle Johnson</i>	229
Building statives in Nsyilxcn		
	<i>John Lyon</i>	237
A corpus-based study of Gitksan modals		
	<i>Lisa Matthewson</i>	251
The definiteness of manners and reasons		
	<i>Marcin Morzycki</i>	269
Crosslinguistic variation in concessive scalar particles		
	<i>Kimiko Nakanishi</i>	279
Sentential negation in Brazilian Portuguese		
	<i>Michel Navarro</i>	293

Ri-investigating inverse number in Dagaare <i>Douglas Pulleyblank, Darius Adjong, Alexander Angsongna, Lisa Matthewson, and Ariwan Addy Suhairi</i>	309
Tell you what. English has quexistentials. <i>Daniel K. E. Reisinger</i>	327
Imperatives and prohibitives in Biblical Hebrew <i>Elizabeth Ritter</i>	343
Ktunaxa ranges <i>Starr Sandoval</i>	353
No vacuous negation in subjunctive questions in Serbian <i>Neda Todorović</i>	365
A maximality – exhaustivity connection: The semantics and pragmatics of <i>-dake</i> in Japanese <i>Satoshi Tomioka</i>	379
The meaning of <i>what</i> <i>Martina Wiltschko</i>	393
Modal adverbs, conjunction reduction, and the structure of coordination <i>Jan-Wouter Zwart</i>	405

Preface

M. RYAN BOCHNAK, EVA CSIPAK, LISA MATTHEWSON,
MARCIN MORZYCKI, AND DANIEL K. E. REISINGER
University of British Columbia

Nobody's perfect, they say.¹ But Hotze nearly is; an alternative title we nearly chose for this volume was *Somebody's Perfect: Papers in Honour of Hotze Rullmann*. In particular, Hotze is an exceptional colleague. We are delighted to be able to honour him with this volume, on the occasion of one of his round-number birthdays. It is also clear that we are not the only people who think highly of Hotze; it was easy to find contributors to the volume who hail from all periods of his academic life.

How is Hotze such a great colleague? Well, we can start with the well-known and obvious fact that he's a great semanticist. He's also — unlike some great semanticists — able to explain semantics clearly and with patience.

Next, we can mention his astounding dedication to students of all levels. Hotze is passionate about teaching and about student well-being. He's an enthusiastic and popular advisor of graduate students. Since Hotze arrived at UBC in 2004, he has helped nearly 50 students obtain Master's or Ph.D. degrees by serving on their thesis committees or as their supervisor. Many of his former or current students have contributed papers to this volume.

Since this isn't a reference letter, we are also allowed to talk with abandon about personal qualities that we appreciate. Hotze has integrity and believes in principles. He is generous with his time and ideas, fair, open-minded, supportive, efficient, 100% reliable, and never competitive. (Hey, nobody get any ideas about poaching him! We need him here.)

One of us, Lisa, has been at UBC even longer than Hotze has. Here are a few words about her personal experience of these past nearly 20 years.

¹ Bertrand, Anne, Yurika Aonuki, Sihwei Chen, Henry Davis, Joash Gambarage, Laura Griffin, Marianne Huijsmans, Lisa Matthewson, Daniel Reisinger, **Hotze Rullmann**, Raiane Salles, Michael Schwan, Neda Todorović, Bailey Trotter, and Jozina Vander Klook 2022. Nobody's perfect. *Languages* 7:148.

Hotze, the day you arrived at UBC, my life got better. It's a pleasure and a fun learning experience to write papers with you. It's a pleasure to co-advise students with you; somehow we always seem to be on the same wavelength. It's a pleasure to co-run research projects with you; ditto. Thank you for being a great listener and for offering sensible words when I occasionally need to vent about life. Thank you for being interested in the same big-picture questions as me and for wholeheartedly and expertly supporting the fieldworking faculty and students in our department. Please don't ever retire!

The title we did choose is a nod to the construction that bears Hotze's name: a Rullmann ambiguity.² What is the range of lengths that the papers in this volume are allowed to be? That's for you to figure out, Hotze; longer than 20 words anyway.

² **Rullmann, Hotze.** 1995. Maximality in the semantics of wh-constructions. Ph.D. dissertation, University of Massachusetts, Amherst.

Degree constructions in Gitksan*

YURIKA AONUKI

Massachusetts Institute of Technology

1 Introduction

This paper offers the first in-depth description of degree semantics in Gitksan and highlights notable properties to narrow down the hypothesis space towards a formal analysis. I demonstrate that comparative and superlative meanings share the same morphological forms and that the difference lies in the size of the comparison class. I then investigate the contributions of the morphemes involved in comparative/superlative constructions, *gay* ‘instead’, *k’aa* ‘exceedingly’, and the preposition *a*. These morphemes are optional, meaning that positive forms can be used in contexts in which both comparatives and superlatives would be used in English. I suggest that consideration of alternative(s) aids the comparative/superlative interpretations in the absence of *k’aa* ‘exceedingly’. One exception is the class of minimum-standard predicates, which require *k’aa* for comparative/superlative interpretations. Finally, I describe the interpretations of measure phrases (MPs) and the division of labour

* I was fortunate to take Hotze’s pragmatics and graduate semantics courses as an undergraduate student. The latter in particular gave me a glimpse of what it would be like to pursue semantic research, and my countless visits to his office hours always made me feel increasingly excited and supported to apply to graduate programs. I was delighted to have him on my MA thesis committee, and I am grateful for his engaging feedback on every chapter and conference abstract, as well as his advice and support beyond my thesis project, which continue to this day.

I would like to thank Gitksan speakers Vincent Gogag and Hector Hill for educating me about the language with much patience and kindness. I thank the Gitksan Lab, especially Michael Schwan, Henry Davis, and Lisa Matthewson, for their support for fieldwork and feedback, as well as Ryan Bochnak, Vera Hohaus, Martin Hackl, Amir Anvari, and the audiences at the BU-Harvard-MIT Summer Fieldwork Symposium and the UBC Linguistics Outside the Classroom for feedback.

between gradable adjectives and nominals, demonstrating that MPs receive differential readings with the former and absolute interpretations with the latter. I stay agnostic about the issue of whether Gitksan should receive a degreeful (Cresswell 1976; von Stechow 1984) or degreeless (Klein 1980) analysis (see Beck, Krasikova, Fleischer, Gergel, Hofstetter, Savelsberg, Vanderelst, and Villalta 2009; Hohaus and Bochnak 2020 for an overview of the cross-linguistic picture) and whether such a binary view of degreefulness is on the right track (Bochnak, Bowler, Hanink, and Koontz-Garboden 2020).

After a brief background on the language (Section 2), Section 3.1 describes the basic pattern of comparative/superlative constructions. Section 3.2 probes for the semantic contributions of the morphemes involved in these constructions. Section 3.3 demonstrates distinct distributions of *k'aa* 'exceedingly' in combination with minimum-standard predicates. Section 3.4 shows differential interpretations of measure phrases occurring with positive and *gay k'aa* constructions. Finally, Section 4 describes other degree constructions in the language, which are incompatible with adjectives and instead involve morphologically related gradable nominals.

2 Language background and previous literature

Gitksan is an Indigenous language spoken in northern British Columbia, Canada. It belongs to the Tsimshianic language family, constituting the Interior Tsimshianic branch along with a neighbouring language, Nisga'a. There are approximately 255 fluent speakers (Gessner, Herbert, and Parker 2022). Unless otherwise noted, the data are from fieldwork with two speakers, Vincent Gogag and Hector Hill.

There has been no formal work on degree constructions in the language family, although some of them have been previously documented. Bicevskis, Davis, and Matthewson (2017) describe amount comparatives and equatives (346–7); Rigsby's (1986) grammar includes degree questions (95–96); and Tarpent's (1987) grammar of Nisga'a documents gradable nominals (244–6) and some comparative sentences (232, 306).

3 Comparative and superlative constructions

3.1 Basic pattern

Comparatives are constructed with the optional morphemes *gay* ‘instead’ and *k’aa* ‘exceedingly’ (1). The standard of comparison is optionally marked by a preposition *a* (1b).¹

- (1) Context: Two children, John and Mary, are standing back to back because they want you to decide who is taller of the two.
- a. Naa=hl (*gay*) (k’aa) ’wii ’nagw-it?
 who=CN instead exceedingly big long-SX
 ‘Who is taller?’ (VG-v., HH-v.)²
- b. (*Gay*) (k’aa) ’wii ’nakw=t³ Mary a[-t]=s
 instead exceedingly big long=PN Mary
 John).
 PREP[-3SG.II]=PN
 ‘Mary is taller (than John).’ (VG, HH-v.)⁴

¹ Glosses follow the conventions in Rigsby (1986). ASSOC: associative; AX: agent extraction; ATTR: attributive; CCNJ: clausal conjunction; CN: common noun connective; COMP: complementizer; DEM: demonstrative; DIST: distal; LVB: light verb; OBL: oblique; PN: proper noun connective; PREP: preposition; PROX: proximal; SX: subject extraction; T: T-morpheme; TR: transitive; Q: question; QUDD: question under discussion downdate; WH: general purpose WH-word; I: series I clitic; II: series II suffix; III series III independent pronouns.

Initials on the right of each example identify the speaker(s) who provided the judgements. “-v.” indicates that the sentence was volunteered by the speaker.

² In (1b), HH volunteered (*Gay*) k’aa ’wii ’nakw=t Mary a=s John. The rest were checked and accepted by both speakers, except that the PP *a=s John* was sometimes omitted, and *Gay* ’wii ’nakw=t Mary was degraded for VG, with a comment that it is acceptable “only if you noted that they are both tall.”

³ Readers may suspect that the combination of the two predicates ’wii ‘big’ and ’nakw ‘long’ may have the effect of intensification. That does not seem to be the case. In the context of describing height, neither ’wii or ’nakw can be used alone. HH rejects both #’Wii=t Michael and #’Nakw=t Michael as a translation of ‘Michael is tall’, remarking that the former is for being large both vertically and horizontally and only used for a baby or child and that the latter would be ‘He’s long.’

Superlatives are expressed by the exact same forms as comparatives (2).

- (2) Context: Trying to stack boxes, with the heaviest one at the bottom.
 Nde=hl *x̣biist* (*gay*) (*k'aa*) *sdiin-it?*
 WH=CN box instead exceedingly heavy-SX
 'Which box is the heaviest?' (VG-v.)

In (3), the same, positive sentence is used in contexts in which both comparative and superlative forms would be used in English, respectively. VG's comment suggests that the felicity of a positive sentence in comparative and superlative contexts is aided by exhaustivity: whichever desk is chosen as the answer to (3), the other salient desk(s) do not count as long in the context. It seems that introduction of alternative(s) is also part of the semantic contribution of *gay* (see section 3.2.1).

- (3) Context: Choosing one desk out of {two, three} desks at a furniture store.
 Guu=hl *ha'niihahle'lst* *'wii 'nagw-it?*
 what=CN desk big long-SX
 'Which desk is long {-er, -est}?' (VG)
 VG (on the comparative context): You're not making a comparison.
 You just want the longest one.

3.2 Probing the contributions of *gay* 'instead', *k'aa* 'exceedingly', and the preposition *a*

In order to investigate the semantic contributions of the three morphemes involved in comparative/superlative constructions, *gay*, *k'aa*, and *a*, and the reasons for their optionality, this section discusses earlier descriptions of these morphemes and provides further data both from within and outside comparatives/superlatives.

⁴ In (1a), VG volunteered the version with *k'aa* only, and HH volunteered the version with *gay k'aa*. The rest were checked and accepted by both speakers.

3.2.1 Gay ‘instead’

Bicevskis et al. (2017) gloss *gay* in comparative constructions as a ‘contrastive’ marker. *Gay* can associate with any lexical element in the sentence and signal that there is a salient alternative to the referent of the associate that makes the proposition false (4).⁵

- (4) a. [Agent] Context: John was supposed to make a cake, but he was too busy, so Mary made it instead.

Gay=t Mary an=t jap[-t]=hl ixsta-m anaax.
 instead=PN Mary AX=1.I make[-2.II]=CN sweet-ATTR bread
 ‘Mary made a cake instead.’ (VG-v.)

- b. [Subject] Context: “Did John sing?”

Nee. Gay=t Mary lim[x]-it.
 no. instead=PN Mary sing-SX
 ‘No, Mary sang instead.’ (VG)

- c. [Object] Context: “Did Mary make fried bread?”

Nee. Gay jab-i-t=hl ixsta-m anaax.
 no. instead make-TR-3.II=CN sweet-ATTR bread
 ‘No, she made a cake instead.’ (HH-v.)

The associate of *gay* is optionally extracted to the sentence initial position following *gay* (4a, 4b). Transitive subject (4a) and intransitive subject (4b) extractions are marked with the overt extraction morphemes *an=* and *-it*, respectively. These extraction patterns are characteristic of A’-dependencies, including *wh*-questions, relative clauses, and focus marking, across the Ts’imshianic family (Aonuki 2022; Brown 2023; Davis

⁵ Rigsby (1986) also documents use of *gay* in *wh*-questions (i).

- (i) Naa an=t gay hlimoo[-t]=s Bruce?
 who AX=3.I gay help[-3.II]=PN Bruce?
 ‘Who (is the one who) helped Bruce?’ (adapted from Rigsby 1986:303)

In addition, Tarpent’s (1987) grammar of Nisga’a documents a morpheme *yagay/yaay/yay* ‘precisely, exactly, instead’, although I have not encountered a use of *gay* in Gitksan as ‘precisely, exactly’. I set aside the question of whether these uses of *gay* correspond to the same morpheme as the one encountered in comparative constructions.

and Brown 2011; Rigsby 1986).

There are at least some indications that *gay* as used in comparative/superlative constructions is indeed the same lexical item as the one in (4). First, extraction is used in comparatives as well. Extraction analogous to (4b), with the associate of comparison immediately following *gay*, is at least sometimes accepted (5) although never volunteered.

- (5) *Gay*=t *Mary* 'wii 'nagw-it.
 instead=PN *Mary* big long-SX
 'Mary is taller.' (VG, HH)

Extraction of the associate of comparison to the sentence-initial position is often volunteered by VG and accepted by HH (6), regardless of whether *gay* is present (6a) or not (6b).

- (6) a. *Hi*'niiluxw-da ii=t *Mary*=hl *gay* 'wii 'nagw-it.
 tall.PL-3PL.INDP CCNJ=PN *Mary*=CN instead big long-SX
 'Both are tall, but Mary is taller.' (VG-v.)
- b. *Mary*=hl 'wii 'nagw-it.
Mary=CN big long-SX
 'Mary is taller.' (VG-v., HH)

Second, HH spontaneously volunteered a *gay* 'instead' sentence with a salient alternative marked with a preposition just like the standard of comparison (7).

- (7) *Gay* 'wii halay=t *Mary* a[-t]=s *John*=aa?
 instead big doctor=PN *Mary* PREP[-3.II]=PN *John*=Q
 'Is Mary a doctor than John?' (volunteered gloss) (HH-v.)
 Elicitor: Is it asking if Mary is more of a doctor than John?
 HH: No, asking if Mary is a doctor and John is not.

3.2.2 *K'aa* 'exceedingly'

K'aa is documented in the grammars of both Gitksan and Nisga'a. Rigsby (1986) glosses *k'aa* in Gitksan as 'exceedingly' (155). Tarpent (1987) glosses *k'aa* in Nisga'a as 'most, excessively, extremely' and describes

that it contributes meanings of intensification, comparatives, or superlatives (389–391). There are some indications that *k'aa* is indeed used for intensification outside of comparatives, at least in VG's dialect.

VG accepts (8) as a translation of 'Michael is very tall.' HH rejects (8), but he accepts *k'aa* in the presence of another modifier *sim* 'truly' or *lukw'il* 'very' (9).

- (8) K'aa 'wii 'nakw=t Michael.
 exceedingly big long=PN Michael
 'Michael is very tall.' (VG:✓,HH:#)

- (9) {Sim, Lukw'il} k'aa 'wii 'nakw=t Michael=is
 truly very exceedingly big long=PN Michael=QUDD
 'Michael is very tall.' (HH)

The sequence of *sim k'aa* has been documented as also modifying a differential degree in Nisga'a (10).

- (10) Sim k'aa 'wiit'axga=t naks-t loo-t.
 really most old=PN spouse-3SG.II OBL-3SG.II
 '{Her, his} {husband, wife} is much older than {her, him}.'
 (Tarpent 1987: 232)

VG volunteered *sim k'aa* as a translation of 'by far' in English, commenting that it is an influence from Nisga'a.

- (11) Sim k'aa 'wii 'nakw=t John.
 very exceedingly big long=PN John
 'John is by far the tallest.' (VG-v.)
 VG: We sometimes borrow from the Ts'imshian proper and from the Naas River people.

As already evident in the data above, the distribution of *k'aa* is quite different from *more* in English. (12) shows that (*gay*) *k'aa* does not have the additive use observed in English *more* (Thomas 2018).

(12) Context: Math question with an answer ‘5’. ...How many berries did she eat?

- a. Gilbil[=hl] maa’y=hl gub-i[-t]=s Mary. Ii=t
 two[=CN] berries=CN eat-TR[-3.II]=PN Mary CCNJ=3.I
 hets’im(h)ux gup[-t]=hl gwila’l[=hl] maa’y.
 again eat[-3.II]=CN three[=CN] berries
 ‘Mary ate two berries. Then she ate three more berries.’
 (VG-v.)

- b. *Gilbil[=hl] maa’y=hl gub-i[-t]=s Mary. Ii=t
 two[=CN] berries-CN eat-TR[-3.II]=PN Mary CCNJ=3.I
 (gay) k’aa gup[-t]=hl gwila’l[=hl] maa’y.
 instead exceedingly eat[-3.II]=CN three[=CN] berries
 intended: ‘Mary ate two berries. Then she ate three more berries.’
 (VG)

3.2.3 *a* is more than a standard marker

I have described in Section 3.1 that the standard of comparison is optionally specified with a preposition *a* (which also surfaces as *e*). Rigsby describes *a~e* as a ‘general preposition’. Strictly speaking, some instances of *a-* in comparatives/superlatives do not mark the standard of comparison. (13) illustrates this.

- (13) a. Naa=hl k’aa ’wii ’nagw-it a=hl
 who=CN exceedingly big long PREP=CN
 sgapdii[-t]=hl tk’ihlxw?
 among[-3.II]=CN children?
 ‘Who is the tallest among the children?’ (VG-v.)

- b. Context: Looking at two people who are both sick.
 Naa=hl k’aa siipxw-it a=hl dip=un?
 who=CN exceedingly sick-SX PREP=CN ASSOC=DEM.PROX
 ‘Who is sicker of these people?’ (HH-v.)

- c. Naa=hl k'aa am-a wil-it a=hl Japan?
 who=CN exceedingly good-ATTR LVB-SX PREP=CN Japan
 'Who is the richest in Japan?' (VG-v.)

The general role of the optional *a~e* phrase in comparatives and superlatives can then be thought of as, instead of only marking the standard of comparison, supplying contextual information that helps narrow down the comparison class, whether that is, e.g., specifying the comparison class directly (13a, 13b), supplying individual(s) that form a comparison class with the associate (i.e., the standard of comparison, (1b)), or specifying the location (13c).

3.3 Minimum-standard gradable adjectives

The data in Section 3.1 showed that *k'aa* 'exceedingly', along with *gay* 'instead', is optional in comparatives and superlatives involving what would be relative adjectives in English. This is not the case with what would be minimum-standard adjectives in English (see Kennedy and McNally 2005; Rotstein and Winter 2004 for classes of gradable adjectives).⁶ While there is variation between speakers and lexical items, one generalization is that, in translations of English comparative sentences with a minimum-standard predicate, *k'aa* is obligatory (14).

(14) Context: Looking at two people who are both sick.

- a. Naa=hl gay k'aa siipxw-it?
 who=CN instead exceedingly sick-SX
 'Who is sicker?' (HH-v.)
- b. Naa=hl k'aa siipxw-it ?(a=hl dip=un)?
 who=CN exceedingly sick-SX PREP=CN ASSOC=DEM.PROX
 'Who is sicker of these people?' (HH)

⁶ I have not found independent diagnostics for minimum-standard predicates. That is, I have not found modifiers like *slightly* and *partially* in English, which are argued to diagnose a minimum-standard predicate (Kennedy and McNally 2005; Rotstein and Winter 2004). For example, *slightly* is translated to Gitksan with *ts'uusxw* 'small'.

- c. Naa=hl (gay) siipxw-it?
 who=CN instead sick-sx
 intended: #‘Who is sicker?’
 ‘Who is sick?’
 HH: You know they’re sick, and you’re asking which one is sick. (HH)

The same pattern is observed in superlatives involving minimum-standard predicates (15).

- (15) Context: Looking at many doors that are all open.
- a. Nde=hl aats’ip (gay) k’aa k’ag-at?
 WH=CN door instead exceedingly door-sx
 ‘Which door is the most open?’ (VG-v.)
- b. Nde=hl aats’ip (gay) k’ag-at?
 WH=CN door instead open-sx
 intended: #‘Which door is the most open?’
 ‘Which door is open?’
 VG: All the rest are closed. (VG)

HH’s and VG’s comments for (14c) and (15b), respectively, are consistent with the view that these predicates have minimum standards that are context-independent (Kennedy 2007): having any degree of sickness or openness would satisfy these predicates regardless of the context.

3.4 Measure phrases

Measure Phrases (MPs) that have been volunteered or recognized by the speakers include *sa* ‘day’ and measurements of length originating from body parts, such as *t’im k’aax* ‘full arm span, fathom’, *hlek moos* ‘inch (lit. crook of thumb)’, and *se’e* ‘foot’.⁷ Use of a whole arm as a measure of length is also reported in Sm’algyax, a.k.a. Coast Tsimshian, as Rigsby (1986:30) reports that *gipl’on* ‘two fathoms’ in Sm’algyax is documented by Dunn (1978). One difference between Sm’algyax and Gitksan is that in the former, MPs using the arm span seem to be suppletive, as in *k’üül k’aay* ‘half fathom’ and *k’ooldq’on* ‘six fathoms’ (First Voices

⁷ See Bicevskis et al. (2017) for MPs used with mass nouns.

2000), while they are formed with a number followed by *t'im k'aax* in the latter. *Hlek moos* was volunteered by VG, and upon being asked whether it is used, HH remarked that the last time he had heard it was in his childhood.

MPs can be the complement of the preposition *a* (16) and serve as the standard.

- (16) Gay k'aa 'wii 'nakw 'nii'y a=hl k'i'y=hl t'im k'aax.
 instead very big long 1SG.III PREP=CN one=CN whole arm
 'I am taller than one arm length.' (VG-v., HH-v.)

MPs are also found sentence initially. Even in combination with morphologically positive constructions without *gay* 'instead' or *k'aa* 'exceedingly', they modify differential rather than absolute degrees (17, 18) (see also Section 4).

- (17) Context: This year is a leap year.
 a. K'i'y=hl sa win 'wii 'nakw[=hl] k'uuhl t=un
 one=CN day COMP big long[=CN] year PN=DEM.PROX
 a=hl gi-k'uuhl.
 PREP=CN last-year
 'This year is one day longer than the last year.' (VG-v.)
 b. (K'am) k'i'y=hl sa win 'wii 'nakw=hl k'uuhl gyuu'n
 only one=CN day COMP big long=CN year now
 a=hl gu-k'uuhl=gi.
 PREP=CN last-year=PR.EVID
 'This year is one day longer than the last year.' (HH-v.)

- (18) K'i'y=hl t'im k'aax win (gay) (k'aa) 'wii
 one=CN whole arm COMP instead exceedingly big
 'nakw[=hl] ha'niitookxw t=un e=s=ust.
 long[=CN] table PN=DEM.PROX PREP=PN=DEM.DIST
 'This table is one arm length longer than that one.' (VG)

It is not clear to me whether the MPs are base-generated in this position and acting as a predicate, taking the clause marked with *win* as

their argument, or they are extracted out of that clause. The morphology is consistent with that of adjunct extraction, which is accompanied by a complementizer and no other morphological marking (Davis and Brown 2011), but there is no instance of a differential MP appearing in its potential base position under such an analysis.

This pattern of MPs receiving differential interpretations in positive constrictions is shared with Japanese (e.g., Aonuki 2023; Kubota 2008; Oda 2008; Sawada and Grano 2011). Moreover, the existence of differential MPs has been argued to be a reliable diagnostic of degreefulness (Deal and Hohaus 2019; von Stechow 1984; cf. Bochnak et al. 2020).

4 Other degree constructions and use of gradable nominals

This section describes some constructions that require gradable nominals rather than adjectives, namely absolute MPs, degree questions, degree demonstratives, and equatives. One exception is that absolute MP readings are possible with minimum-standard adjectives.

Translation of an absolute MP sentence with a relative adjective in English seems to require the prefix *ga-*, at least for VG (19).⁸ Rigsby (1986) describes that *ga-* in Gitksan “forms abstract nominals that signify some attribute or entity” (95), and Tarpent (1987) similarly describes that *ga-* in Nisga’a attaches to an adjectival predicate and forms “an abstract noun” (244).⁹ The entity always follows the *ga-*nominal, intervened by a connective.

⁸ HH pluralizes ‘foot’ to *se-se’e* when the measure is over 1 foot.

⁹ I will refer to *ga-* forms as ‘nominals’, and while there isn’t independent evidence to confirm their nominal category, its distribution described in this section suggests that this assumption is correct.

(21) a. Nde=hl *(ga)-'nagw-i-n?
 WH=CN ga-long-?-2SG.II
 'How tall are you?' (HH-v.)
 (Lit. Where is your height?)

b. *Nde=hl 'nakw 'nit?
 WH=CN long 3SG.III
 'How tall is she?' (VG)

Translations of degree demonstratives (22) and equatives (23) similarly involve a *ga-* nominal.¹¹

(22) a. T=un=hl ga-'nagw-i-t.
 PN=DEM.PROX=CN ga-long-?-3.II
 'She is this tall.' (VG-v., HH-v.)

b. *T=un=hl ('wii) 'nakw 'nit.
 PN=DEM.PROX=CN (big) long 3SG.III
 intended: 'She is this tall.' (VG)

(23) Sagay k'i'y=hl ga-hi'niiluxw-si'm.
 together one=CN ga-tall.PL-1PL.II
 lit. 'Our heights are the same.' (VG-v.)

5 Conclusion

This paper offered the first in-depth descriptions of comparative/superlative constructions and other degree constructions in Gitksan. Comparative/superlative interpretations are available for positive constructions as well as with the optional morphemes *gay* 'instead', *k'aa* 'exceedingly', and

¹¹ The status of the vowel *i* in (21a) and (22a) is unclear at this point. Tarpent (1987) treats its counterpart in Nisga'a as a suffix that occurs with *ga-* (244-246). While in (21a) alone, it could be an epenthetic vowel to break up the sequence of the consonantal sequence in *'nakw-n, that hypothesis would not hold against (22a), where 'nakw-t would be phonologically licit (Henry Davis, Michael Schwan, p.c.).

the preposition *a*. Demonstrating the behaviours of *gay* as an alternative-sensitive morpheme outside comparatives/superlatives, I suggested that positive and *gay* constructions achieve comparative/superlative meanings by consideration of alternative(s), which is contributed pragmatically in the former and semantically in the latter. MPs receive obligatory differential interpretations with gradable adjectives, with the exception of minimum-standard adjectives. Absolute MPs, as well as degree questions, demonstratives, and equatives, require gradable nominals. Questions towards a formal analysis include 1) what the semantic contribution of *k'aa* is such that it is obligatory for comparatives/superlatives with minimum-standard adjectives but not relative adjectives; 2) why MPs receive differential interpretations with relative adjectives; 3) what the source(s) of the distinct behaviours of minimum-standard adjectives is formally; and 4) what explains the division of labour between gradable adjectives and nominals across degree constructions.

References

- Aonuki, Yurika. 2022. Free relatives in Gitksan. In *Proceedings of Semantics of Under-Represented Languages in the Americas (SULA) 11*.
- Aonuki, Yurika. 2023. Evaluativity in Japanese degree constructions. Handout from *Sinn und Bedeutung 28*.
- Beck, Sigrid, Sveta Krasikova, Daniel Fleischer, Remus Gergel, Stefan Hofstetter, Christiane Savelsberg, John Vanderelst, and Elisabeth Villalta. 2009. Crosslinguistic variation in comparison constructions. *Linguistic Variation Yearbook* 9:1–66.
- Bicevskis, Katie, Henry Davis, and Lisa Matthewson. 2017. Quantification in Gitksan. In *Handbook of quantifiers in natural language: Volume ii*, 281–382. Springer.
- Bochnak, M Ryan, Margit Bowler, Emily A Hanink, and Andrew Koontz-Garboden. 2020. Degreefulness is the result of functional inventory, not a parameter. Handout from *Sinn und Bedeutung 25*.
- Brown, Collin. 2023. Questions and their relatives in Sm'algyax. Ms., UCLA. To appear in *International Journal of American Linguistics*.

- Cresswell, Max J. 1976. The semantics of degree. In *Montague grammar*, 261–292. Elsevier.
- Davis, Henry, and Jason Brown. 2011. On A'-dependencies in Gitksan. In *Proceedings of the 46th International Conference on Salish and Neighbouring Languages*, 43–80. University of British Columbia Working Papers in Linguistics.
- Deal, Amy Rose, and Vera Hohaus. 2019. Vague predicates, crisp judgments. In *Proceedings of Sinn und Bedeutung 23*, 347–364.
- First Voices. 2000. Sm'algyax words. Retrieved from <https://www.firstvoices.com/explore/FV/sections/Data/Ts'msyen/Sm%E2%80%9999algyax/Sm'algya%CC%B1x/learn/words?page=1&pageSize=10>.
- Gessner, Suzanne, Tracey Herbert, and Aliana Parker. 2022. Report on the status of B.C. First Nations languages 4th edition. First Peoples' Cultural Council. Retrieved from <https://fpcc.ca/stories/status-of-languages>.
- Hohaus, Vera, and M Ryan Bochnak. 2020. The grammar of degree: Gradability across languages. *Annual Review of Linguistics* 6:235–259.
- Kennedy, Christopher. 2007. Modes of comparison. In *Proceedings from the Annual Meeting of the Chicago Linguistic Society*, volume 43, 141–165. Chicago Linguistic Society.
- Kennedy, Christopher, and Louise McNally. 2005. Scale structure and the semantic typology of gradable predicates. *Language* 81:1–37.
- Klein, Ewan. 1980. A semantics for positive and comparative adjectives. *Linguistics and Philosophy* 4:1–45.
- Kubota, Yusuke. 2008. Phrasal comparatives in Japanese: A measure function-based analysis. *Empirical Issues in Syntax and Semantics* 7:1–20.
- Oda, Toshiko. 2008. *Degree constructions in Japanese*. Doctoral dissertation, University of Connecticut.
- Rigsby, Bruce. 1986. Gitksan grammar. Manuscript, University of Queensland.
- Rotstein, Carmen, and Yoad Winter. 2004. Total adjectives vs. partial

- adjectives: Scale structure and higher-order modifiers. *Natural Language Semantics* 12:259–288.
- Sawada, Osamu, and Thomas Grano. 2011. Scale structure, coercion, and the interpretation of measure phrases in Japanese. *Natural Language Semantics* 19:191–226.
- von Stechow, Arnim. 1984. Comparing semantic theories of comparison. *Journal of Semantics* 3:1–77.
- Tarpent, Marie-Lucie. 1987. A grammar of the Nisgha language. Doctoral dissertation, University of Victoria.
- Thomas, Guillaume. 2018. Underspecification in degree operators. *Journal of Semantics* 35:43–93.

Incorporating sign language phonetics & phonology exercises into the linguistics classroom*

YURIKA AONUKI

Massachusetts Institute of Technology

KATHLEEN CURRIE HALL

University of British Columbia

1 Introduction

In many introductory linguistics courses, instructors make a point of including at least a brief section on sign languages, trying to help students understand that languages exist in multiple modalities¹ and have analyzable linguistic structures. Often, though, after that first mention, sign languages essentially disappear from the standard linguistics curriculum, potentially leaving students with an impression that linguistic research on sign languages is not possible or not active, and/or that such research is somehow secondary to ‘main-stream’ research on spoken languages. We

* We gratefully acknowledge the support and influence of the various people who have helped shape our perspectives on the topics discussed here. First and foremost, this includes our own ASL instructors, who have been generous and patient with their knowledge sharing. We are also grateful to our students who have accompanied us in our evolving journey, to Maya Honda for encouraging us to write about our experiences, and to the editors of this volume for giving us a platform to do so. We also want to acknowledge and express our gratitude for the passion and dedication Hotze Rullmann has shown for undergraduate education. His time in the classroom and years as undergraduate advisor continue to make a lasting positive impact on both our lives and those of our students, and we hope this chapter helps to commemorate his belief in the importance of excellent education.

The author order is alphabetical.

¹ Our discussion in this paper focuses only on visual sign languages, but we acknowledge the existence of tactile sign language(s) as well; see e.g. granda and Nuccio (nd) and Edwards and Brentari (2020) for discussion of protactile sign in particular. We have not yet included these in our classroom discussions and exercises, but hope to do so in the future.

doubt that most instructors intend for this to be the lesson learned, and recognize that changing the standard practice can be time consuming and difficult. One difficulty is that, while there are many linguistic similarities across the two modalities, there are also important differences, which instructors may not be familiar with and/or may feel are too complicated to ‘get into’ when there is limited time in a course. This is not a new observation; others have discussed this problem as well (see, e.g., Hochgesang 2019; Lillo-Martin and Hochgesang 2022; Sanders, Umbal, and Konnelly 2020; Zuraw 2022).

Our goals in this paper are (1) to motivate the importance of including sign languages in the linguistics classroom, from sociocultural, empirical, and theoretical perspectives, and (2) to share our experiences of getting started on such efforts in our phonetics and phonology courses, from a practical perspective.

It is important for us to acknowledge that we are both hearing and in the process of learning a sign language, specifically American Sign Language (ASL), and about Deaf culture(s). We hope that, by foregrounding d/Deaf-authored² resources that we have consulted, we contribute to ensuring that future discussions of sign languages in the classroom are accurate and culturally sensitive. We also hope that this paper addresses some of the concerns that may have been keeping other hearing instructors from taking the first step in starting such discussions.

The paper is structured as follows: we start by sharing our motivation for incorporating sign language data throughout the curriculum (Section 2). This is followed by a discussion of some of the practical considerations that we have faced in the process of including sign languages in our courses (Section 3), namely understanding the sociocultural contexts of sign languages (Section 3.1), thinking about finding data sources (Section 3.2), and dealing with the question of how to present data to students (Section 3.3). This includes the issue of transcription, which we see as one of the major obstacles in engaging students in in-depth phonetic and

² We follow the convention to “use the lowercase *deaf* when referring to the audiological condition of not hearing, and the uppercase *Deaf* when referring to a particular group of deaf people who share a [sign] language...and a culture” (Padden and Humphries 1988: 2). See also discussion in e.g. Padden and Humphries 2005, who note that the convention was started by James Woodward in the 1970s.

phonological analyses of sign languages within the limited course time. These topics are illustrated by exercises that we have designed for use in the undergraduate curriculum.

2 Motivation for incorporating sign language data

2.1 Sociocultural and empirical perspectives

The linguistics classroom can and should be a place where students can become increasingly aware of minority communities through respectful and accurate discussions of minority and/or under-represented languages. In the context of sign languages, our goal is to ensure that every student coming out of our curriculum has accurate knowledge of sign languages, Deaf cultures, and issues of audism (“the notion that one is superior based on one’s ability to hear or behave in the manner of one who hears,” coined and defined in Humphries 1977: 11–12), so that they can be allies to d/Deaf communities in the larger society.

As Henner and Robinson (2023) point out, *modality chauvinism*, which they define as “beliefs and actions that support the superiority of one modality over others” (11), is still perpetuated in research and teaching practices in theoretical linguistics, both through limited discussions of sign languages and the nature of such discussions. One of the examples they provide is the common practice of defining *linguistics* (and phonetics and phonology in particular) with focus on speech sounds and no reference to sign languages. They remark, “[t]he artificial limitation of linguistics to speech is an extension of the cultural belief that the most or only valid languaging is speech” (12). It seems then that the linguistics classroom must be one of the first places to challenge such a belief. Of course, this presupposes that we as researchers and instructors are aware of the issues and educate ourselves (see also Section 3.1). Lillo-Martin and Hochgesang (2022) similarly emphasize the importance of instructors who are not involved in research on sign languages nonetheless understanding such research and including discussions of sign languages in their teaching. They argue that failure to include sign languages in linguistic theories is failure to understand the full capacity of the language faculty.

At the same time, we are not advocating for having sign languages be used in examples for *all* topics in all courses. However, we do think that

it is essential to incorporate discussions of sign languages into the curriculum in a periodical, programmatic manner rather than consolidating them into a section specifically on sign languages, or relegating them to a separate course. Again, the latter approaches would likely reinforce the misconception that research on signed languages is somehow secondary to that on spoken languages.

2.2 Theoretical perspectives

From the perspective of theoretical training, one significant merit of discussing sign language data in the classroom is that actively comparing how the same linguistic properties manifest in both spoken and signed languages helps reinforce students' understanding of theoretical concepts they encounter. For example, in an introductory phonology course after the concept of minimal pairs has been introduced, a data set from a sign language can provide an opportunity both for the instructor to check the students' understanding of the concept and for the students to practice applying the concept. Figure 1 illustrates such an exercise. A set of words from ASL are provided, with some of them forming minimal pairs or near minimal pairs with each other (the figure shows an excerpt from the original, which had nine signs for students to examine).³

This is a new challenge for the students; through similar minimal pair exercises on spoken languages, they may have developed a habit of looking for two strings of IPA transcriptions that differ in one symbol, but the same 'recipe' would not work when faced with raw data from a sign language. Without a transcription, and especially if they have limited knowledge of each parameter in the signed modality, the students may have no preconceptions about which parameters they should argue to be contrastive. They must extend their understanding of the concept of contrast and a 'minimal pair' to a more abstract level. At the same time, working on this exercise allows the students to find out about the highly

³ We recognize that the use of raw, visual data instead of machine-legible transcriptions in the exercises presented in this paper may make them inaccessible to blind and deafblind students and instructors. This issue is partially due to the lack of universally adopted, machine-legible transcription systems (see Section 3.3; though see SiGML (Elliott, Glauert, Jennings, and Kennaway 2004)). In this particular case, the issue could be addressed by providing prose descriptions of each element in each video.

Parameters in ASL

Below are data from American Sign Language (ASL). Among handshape, location, movement, orientation, the number of the hands involved, and duration, which ones can be argued to be contrastive in ASL, and why? Use the data to support your answer. (The images are the final state of the sign. Click on the link to see each video.)



Figure 1: An example of a minimal pair exercise in ASL, with data from Sehyr et al. (2021).

simultaneous nature of phonological structures in sign languages, even with relatively little theoretical knowledge of sign language phonology. See also Figure 3 in Section 3.3 for an exercise using a morphological process in ASL to reinforce the concept of auto-segmental representation.

3 Practical considerations

Once the decision has been made to incorporate sign languages into the linguistics curriculum, the actual implementation can begin. There are a variety of considerations that make doing research with signed language data different from doing research with spoken language data (see e.g., Quer and Steinbach 2019), and the same considerations hold when trying to incorporate these languages into our classrooms. In this section, we focus on three issues: understanding the sociocultural context of sign languages, obtaining linguistic data, and presenting those data to students.

3.1 Sociocultural contexts

The very first step is to learn about the cultural contexts of sign languages.⁴ One thing to keep in mind is that some of the materials about sign languages found online, including language teaching materials, are created by hearing people who are not fluent in the language nor part of a Deaf community. Use of inauthentic materials by hearing people would contribute to misrepresentation and cultural exploitation, and we recognize that ignorance on our part could have a negative sociocultural impact on d/Deaf communities. We check the resources we use for the cultural and linguistic status of the authors, in addition to the licensing and permissions associated with the re-use of such materials.

This is not to say that hearing people can't be fluent signers nor that they can't research sign languages. Nevertheless, we do prioritize highlighting the works of d/Deaf signers whenever possible.⁵ Our own literacy for finding culturally authentic sources primarily comes from learning ASL from Deaf instructors and meeting members of the local Deaf and signing communities, and we are in a continuing learning process. We believe that cultural sensitivity can be fostered in the classroom as well. In fact, discussions about finding linguistically and culturally accurate resources about sign languages periodically come up in our classroom, especially in upper-year seminars with a research paper component.

Some examples that we have used for educating both ourselves and our students include the book *Inside Deaf Culture* (Padden and Humphries

⁴ We think it's useful for linguistics instructors to think about the cultural context of *all* languages they include in their courses, but also recognize that it's not practical to go into depth if one includes a diverse range of languages. However, for sign languages in particular, there is a wealth of misinformation that students often come in with, and we do think it is imperative that some of it be addressed (especially for programs in which students are likely to be interested in careers in Speech Language Pathology or Audiology).

⁵ Note that we specifically avoid referencing 'native' signers, regardless of hearing status. See e.g. Cheng, Burgess, Vernooij, Solís-Barrosol, McDermott, and Namboodiripad (2021) for discussion of why the concept of a 'native' user of a language is problematic for all languages and Quer and Steinbach (2019) for discussion of why it's particularly problematic for sign languages. See also Hener and Robinson (2023) for more general discussion of the problematic history in linguistics of assigning value judgments to different ways of using language, including the idea of 'fluency.'

2005) by Carol Padden and Tom Humphries, who are long-time leaders in the Deaf ASL language / linguistics studies community, and the novel *True Biz* by Deaf writer and instructor Sara Nović (Nović 2022), which introduces a lot of context for Deaf culture and the importance of access to sign language in a less formal/academic style. It is also important to note that many sources authored by members of Deaf communities will themselves be in sign languages and therefore video-based: the documentary *Audism Unveiled* (Bahan, Bauman, and Montenegro 2008) is produced by members of the Deaf Studies department at Gallaudet University and entirely narrated in ASL (with subtitles in several spoken languages), and there are also various YouTube playlists of presentations hosted by e.g. TEDx Gallaudet (TED 2014). The website HandSpeak® (Lapiak 1995), created and maintained by Deaf and natively signing ASL instructor and literary media creative Jolanta Lapiak, offers a collection of her articles on ASL and Deaf culture that are searchable by topic. This is by no means an exhaustive list of resources — just a starting point for instructors looking for Deaf-led general introductory materials.

We should say, too, that starting the initial discussion of sign languages in an introductory course involves devoting designated class time for both cultural and theoretical contextualization. Devoting this time in an introductory course means that students acquire the basic foundation that instructors of subsequent courses can build on whenever data from a sign language become relevant. Otherwise, having sign languages appear only sporadically in a student’s curriculum can lead to additional practical difficulties in re-introducing foundational concepts across courses.

3.2 Sources of linguistic data

Once the sociocultural foundations are laid, specific sign language data for illustrating particular linguistic concepts need to be found. Given both the visual nature of sign languages and the fact that many are understudied, we are often required to look for sources outside of the traditional academy (see also Quer and Steinbach 2019). Hou, Lepic, and Wilkinson (2022) discuss many of the practical and ethical considerations that arise when using sign language data collected from the internet for research purposes, and again, many of the same points hold for developing data sets for the classroom. One particular consideration is that instructors

may have a limited understanding of dialectal differences within a sign language. To avoid unknowingly misrepresenting data from multiple dialects as a single dialect, whenever possible, we try to use data signed by a single person within one exercise, unless we have a good reason to think that the data come from the same variety (or are specifically trying to illustrate linguistic variation).

For ASL in particular, most of our existing exercises draw from five lexical databases / dictionaries: (1) the dictionary on HandSpeak® (Lapiak 1995; <https://www.handspeak.com/>); (2) the dictionary on a site called ‘American Sign Language University,’ created by Deaf professor Bill Vicars (Vicars 1997; <http://www.lifepprint.com/index.htm>); (3) the online lexical database ASL-Lex (Sehyr et al. 2021; <https://asl-lex.org/index.html>); (4) the ASL portion of the global SignBank project (<https://signbank.cls.ru.nl/>), ASL SignBank (Hochgesang, Crasborn, and Lillo-Martin 2023; <https://aslsignbank.haskins.yale.edu/>); and (5) the print-only *Canadian Dictionary of American Sign Language* (Bailey and Dolby 2002).

One excellent starting place for looking for sign language data in sign languages beyond ASL is the ‘Sign Language Dataset Compendium’ (Kopf, Schulder, and Hanke 2022; <https://www.sign-lang.uni-hamburg.de/lr/compendium/index.html>), “an overview of digital resources for signed languages suitable for research.” It includes resources for more than 80 different sign languages, including both corpora and lexical / dictionary resources, and provides information on what data are available and how they may be accessed, shared, and cited.⁶

Another resource is Berez-Kroeker, McDonnell, Koller, and Collister (2022), which contains several chapters specifically dealing with different kinds of sign language data. The focus of the volume is on data management for research use, but there are many references to existing data sources that may prove useful for instructors looking for data sets. Relevant chapters include Palfreyman (2022) on fieldwork data, Hochgesang (2022) on acquisition data, and Crasborn (2022) on corpora. Relatedly, Fenlon and Hochgesang 2022 is an entire volume on sign language corpora, with a dedicated chapter on utilizing such sources (Börstell 2022).

⁶ Note that being listed in the compendium does not mean that the data are freely available for use in exercises or publications; each source simply has its license information listed.

3.3 Presenting data to students

The final practical consideration we address here is how to actually present sign language data to students, with specific focus on the phonetic and phonological domain. Many of the differences between signed and spoken languages come down to the very difference in modality (see e.g. Meier 2002; Quer and Steinbach 2019). That is, while there are many similarities in phonological structure between signed and spoken languages, the fact that they are communicated using different modes results in many apparent differences. For example, signed and spoken languages are similar in that they both show duality of patterning; they both have ‘phonemic’ elements that can be substituted to form minimal pairs; these phonemic elements can be broken down into phonological features; these various elements have hierarchical structure; there are processes like assimilation and deletion that can apply to these elements; there are higher-level, prosodic elements; and phenomena like markedness govern the distribution of these elements (see e.g., Fenlon, Cormier, and Brentari 2015; Sandler 2012). However, the details of each of these areas diverge when it comes to actual implementation (oral/acoustic vs. corporeal/visual). While in some sense, this difference in the substance of the elements is ‘small’, it is also *fundamental* when every structural similarity to be found in the domain of phonology is embedded in, and perhaps masked by, the physical substance. In a classroom, especially in introductory courses, students are often only beginning to understand how to think about *any* language in terms of its internal structure. Such unfamiliarity with the basic elements then magnifies the apparent differences between signed and spoken languages, and these differences are not at all insignificant when it comes to including exercises on sign languages in the classroom.

Particularly notable is the fact that language data in these two modalities must be represented differently. Much of the data typically included in linguistic exercises is transcribed, and for spoken languages, this usually means using the International Phonetic Alphabet (IPA) or an adaptation of it. Teaching students about the IPA and having them practice it enough to at least recognize transcriptions generally takes at least a week in our introductory courses, and often focuses on transcription of a language that the students are presumed to be familiar with (e.g., English at

universities in Anglo Canada), with the expectation that students can then extrapolate the principles to other spoken languages as they encounter such data. Because sign languages use a different modality, the transcription system *cannot* be the same. This means that, if transcribed data are to be used, a significant period of time would also be needed to teach students about a second method of transcription, and the amount of time needed would likely be even greater than the time it takes to introduce students to IPA, because of the greater degree of starting unfamiliarity with sign languages for many students. Added to this are the facts that there is no single agreed-upon transcription system for signed languages akin to the widespread acceptance of the IPA for spoken languages and that many instructors are themselves not already familiar with such systems as do exist (e.g., Stokoe notation (Stokoe, Casterline, and Croneberg 1965), Prosodic Model notation (Eccarius and Brentari 2008), the Hamburg Notation System (HamNoSys; Prillwitz, Leven, Zienert, Hanke, and Henning 1987), or Sign Language Phonetic Annotation (Johnson and Liddell 2010, 2011a,b, 2012, 2021; Liddell and Johnson 2019); for discussion of these various systems, see Hochgesang 2014). This means that choosing a means of representation to make sign language data accessible is not trivial.

In many cases, it may be easiest to simply represent the data in visual form, as images or videos. While this can be effective, it should also be approached with caution. There are reasons that instructors tend to present data in phonologically transcribed forms to students: such forms have already been ‘massaged’ by the transcriber to reflect the important information and level of detail needed for further analysis. Most instructors would probably consider handing introductory students a set of sound files from an unfamiliar language and asking them to analyse a phonological process within them a completely ridiculous idea. The implausibility of such a task largely comes from the level of detail present in a recording; how is a student to know from a small data set whether, e.g., variations in pitch are phonemic, syntactic, semantic/pragmatic, or accidental in nature? The same is true for sign language data: unless the viewer is familiar with the language, there is no way to know which formational elements play which role. For example, consider the two signs shown in Figure 2. These are still images from tokens of two different lexical items in ASL; the one on the left is an image from the sign for the third person singu-

lar reflexive pronoun (HIMSELF/HERSELF), while the one on the right is an image from the sign for BEAT or ABUSE. Which elements are phonologically important? In fact, the key difference between these signs as visible here⁷ is the extension of the thumb on the right hand in the reflexive, as compared to its being folded under in BEAT.



Figure 2: Two different lexical items in ASL, SELF on the left (Vicars 1997; <http://www.lifeprint.com/asl101/pages-signs/s/self.htm>) and BEAT on the right (Vicars 1997; https://www.youtube.com/watch?v=Y5NQ0WsJ_zk), illustrating both phonological differences and non-lexical variation.

At the same time, there are visible differences in other properties of these tokens. First, this token of BEAT has a facial expression absent in SELF (3.sg). Second, the angle at which the left index finger is pointing is different between the two pictures. While both facial expression and hand orientation can be lexically contrastive, they are not in this particular instance. (For evidence of the non-lexical status of these properties in these signs, compare the tokens pictured here to those shown on Handspeak[®] (Lapiak 1995), for example: the second video at <https://www.handspeak.com/word/3584/> (SELF, 3.sg.) vs. <https://www.handspeak.com/word/5096/> (BEAT).)⁸

⁷ The movements are also different in the signs, but that is not detectable in a still image without adding e.g. arrows.

⁸ That is not to say that the orientation and facial expression are not related to the semantic content of the sign. This might be considered somewhat similar to the fact that the English word ‘beat’ is lexically /bit/ but could be produced with different pitch, volume, rhythm, or voice quality characteristics, some of which

The point here is that examining ‘raw’ data, while it avoids the problem of transcription, also deprives students of the benefits of clean transcriptions to compare. This problem of representation has cascading effects for instructors: almost all of our typical phonological exercises involve transcribed data; how do we present an advanced problem on a specific topic if there’s not a foundational transcription system to use? More broadly, given the limited amount of time to discuss these issues in most courses, how can we set students up for a successful understanding of linguistic structure with sign language examples, when there are so many modality-specific representational differences?⁹

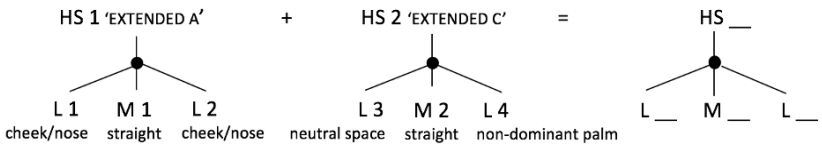
The approach we have used most has been scaffolding, or providing students with basic support for tasks that are somewhat outside of their current state of understanding (Wood, Bruner, and Ross 1976), when discussing sign language data at any point of the curriculum. As an example, consider an exercise from a relatively advanced phonology course, after the concept of auto-segmental representation is introduced. Data about phonological assimilation in ASL compounds provide an excellent opportunity for testing and reinforcing students’ understanding. An excerpt from one such exercise is in Figure 3. This exercise requires students to figure out which phonological elements of each component sign end up in the compound sign. In doing so, the students are expected to observe and describe instances of delinking and spreading. However, other than having a basic background in knowing that sign languages have linguistic structure, students do not need to be able to phonetically analyze the data or read transcriptions. Instead, all relevant vocabulary is included in the exercise itself, and the actual response from students involves simply giving the labelling number for the handshape, movement, and locations that appear in the final compound; no transcription system is needed.

may illustrate or reflect a particular instance of beating.

⁹ Also important, though not addressed here, is that instructors need to be careful in interpreting data. It is important for instructors who are not themselves familiar with sign language data that they are using not to make assumptions about the phonological structure, but rather to rely on expert analysis.

Compounds in ASL

Below is a sample representation of a compound sign WIFE using the Hand-Tier model (Sandler and Lillo-Martin 2006). “L” stands for Location and is defined as “the starting and ending point that the hand traverses in articulating the sign” (p. 133). “M” stands for Movement. “HS” is a shorthand for handshape. Complete the representation of the compound sign WIFE. Fill in the blanks with numbers to indicate which HS, L, and M from the component signs are seen in the compound.



GIRL: Thumbtip of right EXTENDED A hand, palm left, is placed on the right cheek and is stroked forward/downward. Motion may be repeated.



MARRY: EXTENDED C hands are held apart with palms facing and the right hand slightly above the left. The hands are then clasped together.



WIFE: Thumbtip of right EXTENDED C hand, palm down, is positioned at right cheek and the hand is brought downward to clasp left EXTENDED C hand, of which palm faces upward.

Figure 3: An example of an exercise on compounding in ASL, with data from Bailey and Dolby (2002: 279, 406, 825).

Similar approaches can be used for almost any level of exercise. The example in Figure 1 above, for example, uses these techniques. The instructions explicitly provide the list of characteristics to focus on, and brief descriptions of each of these could also be provided if needed. Other examples might include having students match signs to descriptions of the phonetic content to learn about articulation or to specific phonetic transcriptions using one (or more!) of the proposed transcription systems for sign languages to learn about transcription, or having them put a set of

signs in order of phonological markedness or predicted age of acquisition, etc. The key idea of scaffolding in this context is that the instructor has done a fair bit of work on the selection of individual signs and provided relevant descriptions, to allow the students to focus on the conceptual structure of an analysis and not worry about the vocabulary and lack of transcription of the specific items.

4 Conclusion

In this paper, we have shared our experiences of developing and incorporating phonetic and phonological exercises on data from sign languages in the linguistics classroom. Our motivation for this effort includes 1) increasing our students' awareness of sign languages as a fundamental part of empirical linguistic data and 2) enhancing our students' understanding of theoretical concepts by applying them to different modalities. We have addressed some practical considerations, namely understanding the sociocultural contexts and prioritizing d/Deaf-led resources in that process, finding linguistic data, and presenting data to students, with particular focus on the issue of transcription.

We hope that our lessons learned, specific pointers to cultural and linguistic resources, and examples of exercises we have developed have addressed some of the initial challenges likely faced by readers considering undertaking similar efforts. We strongly hope that such readers will not stop at our paper. Instead, it is intended as a starting point, for referring to the resources cited here and beyond, especially those authored by d/Deaf people, and for thinking in practical terms about incorporating discussions of sign languages into the curriculum in accurate and culturally sensitive ways.

References

- Bahan, Ben, H-Dirksen Bauman, and Facundo Montenegro. 2008. *Audism unveiled*. San Diego: Dawn Sign Press.
- Bailey, Carole Sue, and Kathy Dolby, ed. 2002. *The Canadian dictionary of ASL*. The University of Alberta Press.
- Berez-Kroeker, Andrea L., Bradley McDonnell, Eve Koller, and Lau-

- ren B. Collister, ed. 2022. *The open handbook of linguistic data management*. MIT Press.
- Börstell, Carl. 2022. Searching and utilizing corpora. In *Signed language corpora*, ed. Jordan Fenlon and Julie A. Hochgesang. Gallaudet University Press.
- Cheng, Laretta S. P., Danielle Burgess, Natasha Vernooij, Cecilia Solís-Barrosol, Ashley McDermott, and Savithry Namboodiripad. 2021. The problematic concept of native speaker in psycholinguistics: Replacing vague and harmful terminology with inclusive and accurate measures. *Frontiers in Psychology* 12. URL <https://www.frontiersin.org/articles/10.3389/fpsyg.2021.715843>.
- Crasborn, Onno. 2022. Managing data in sign language corpora. In *The open handbook of linguistic data management*, ed. Andrea L. Berez-Kroeker, Bradley McDonnell, Eve Koller, and Lauren B. Collister, 463–470. MIT Press. URL <https://doi.org/10.7551/mitpress/12200.001.0001>.
- Eccarius, Petra, and Diane Brentari. 2008. Handshape coding made easier: A theoretically based notation for phonological transcription. *Sign Language and Linguistics* 11:69–101.
- Edwards, Terra, and Diane Brentari. 2020. Feeling phonology: The conventionalization of phonology in protactile communities in the United States. *Language* 96:819–840.
- Elliott, Ralph, John Glauert, Vince Jennings, and Richard Kennaway. 2004. An overview of the SiGML notation and SiGMLSigning software system. *sign-lang@LREC 2004* 98–104.
- Fenlon, Jordan, Kearsy Cormier, and Diane Brentari. 2015. The phonology of sign languages. In *The Routledge handbook of phonological theory*, ed. S. J. Hannahs and Anna Bosch. Routledge.
- Fenlon, Jordan, and Julie A. Hochgesang, ed. 2022. *Signed language corpora*. Gallaudet University Press.
- granda, aj, and Jelica Nuccio. nd. Protactile principles. World Association of Sign Language Interpreters. <https://wasli.org/special-interest/deafblind-interpreting>.

- Henner, Jon, and Octavian Robinson. 2023. Unsettling languages, unruly bodyminds: A Crip Linguistics manifesto. *Journal of Critical Study of Communication and Disability* 1:7–37. URL https://doi.org/10.48516/jescd_2023vol1iss1.4.
- Hochgesang, Julie. 2022. Managing sign language acquisition video data: A personal journey in the organization and representation of signed data. In *The open handbook of linguistic data management*, ed. Andrea L. Berez-Kroeker, Bradley McDonnell, Eve Koller, and Lauren B. Collister, 367–383. MIT Press. URL <https://doi.org/10.7551/mitpress/12200.001.0001>.
- Hochgesang, Julie A. 2014. Using design principles to consider representation of the hand in some notation systems. *Sign Language Studies* 14:488–542.
- Hochgesang, Julie A. 2019. Inclusion of Deaf linguists and signed language linguistics. Georgetown University Roundtable (GURT) 2019. https://figshare.com/articles/conference_contribution/Inclusion_of_Deaf_Linguists_and_Signed_Language_Linguistics/13549316
Film: <https://youtu.be/1tUHS3reLeE>.
- Hochgesang, Julie A., Onno Crasborn, and Diane Lillo-Martin. 2023. ASL Signbank. New Haven, CT: Haskins Lab, Yale University. <https://aslsignbank.haskins.yale.edu/>.
- Hou, Lynn, Ryan Lopic, and Erin Wilkinson. 2022. Managing sign language video data collected from the internet. In *The open handbook of linguistic data management*, ed. Andrea L. Berez-Kroeker, Bradley McDonnell, Eve Koller, and Lauren B. Collister, 471–480. MIT Press. URL <https://doi.org/10.7551/mitpress/12200.001.0001>.
- Humphries, Tom. 1977. Communicating across cultures (deaf-hearing) and language learning. Doctoral Dissertation, Union Graduate School. URL <https://www.proquest.com/dissertations-theses/communicating-across-cultures-deaf-hearing/docview/302878949/se-2?accountid=14656>.
- Johnson, Robert E., and Scott K. Liddell. 2010. Toward a phonetic representation of signs: Sequentiality and contrast. *Sign Language Studies* 11:241–274.

- Johnson, Robert E., and Scott K. Liddell. 2011a. A segmental framework for representing signs phonetically. *Sign Language Studies* 11:408–463.
- Johnson, Robert E., and Scott K. Liddell. 2011b. Toward a phonetic representation of hand configuration: The fingers. *Sign Language Studies* 12:5–45.
- Johnson, Robert E., and Scott K. Liddell. 2012. Toward a phonetic representation of hand configuration: The thumb. *Sign Language Studies* 12:316–333.
- Johnson, Robert E., and Scott K. Liddell. 2021. Toward a phonetic description of hand placement on bearings. *Sign Language Studies* 22:131–180.
- Kopf, Maria, Marc Schulder, and Thomas Hanke. 2022. The Sign Language Dataset Compendium: Creating an overview of digital linguistic resources. In *Proceedings of the LREC2022*, ed. Eleni Efthimiou, Stavroula-Evita Fotinea, Thomas Hanke, Julie A. Hochgesang, Jette Kristoffersen, Johanna Mesch, and Marc Schulder, 102–109. ELRA. URL <https://www.sign-lang.uni-hamburg.de/lrec/pub/22025.pdf>.
- Lapiak, Jolanta. 1995. Handspeak[®]. <https://www.handspeak.com/>.
- Liddell, Scott K., and Robert E. Johnson. 2019. Sign language articulators on phonetic bearings. *Sign Language Studies* 20:132–172.
- Lillo-Martin, Diane, and Julie A Hochgesang. 2022. Signed languages—unique and ordinary: A commentary on Kidd and Garcia (2022). *First Language* 42:789–793.
- Meier, Richard P. 2002. Why different, why the same? Explaining effects and non-effects of modality upon linguistic structure in sign and speech. In *Modality and structure in signed and spoken language*, ed. Richard P. Meier, Kearsy Cormier, and David Quinto-Pozos, 1–26. Cambridge University Press. URL <https://doi.org/10.1017/CBO9780511486777.001>.
- Nović, Sara. 2022. *True biz*. Random House.
- Padden, Carol, and Tom Humphries. 1988. *Deaf in America: Voices from a culture*. Harvard University Press.

- Padden, Carol, and Tom Humphries. 2005. *Inside Deaf culture*. Harvard University Press.
- Palfreyman, Nick. 2022. Managing sign language data from fieldwork. In *The open handbook of linguistic data management*, ed. Andrea L. Berez-Kroeker, Bradley McDonnell, Eve Koller, and Lauren B. Collier, 267–276. MIT Press. URL <https://doi.org/10.7551/mitpress/12200.001.0001>.
- Prillwitz, Sigmund, Regina Leven, Heiko Zienert, Thomas Hanke, and Jan Henning. 1987. *HamNoSys: Hamburg Notation System for sign languages: An introduction*. Zentrum für Deutsche Gebärdensprache.
- Quer, Josep, and Markus Steinbach. 2019. Handling sign language data: The impact of modality. *Frontiers in Psychology* 10. URL <https://doi.org/10.3389/fpsyg.2019.00483>.
- Sanders, Nathan, Pocholo Umbal, and Lex Konnelly. 2020. Methods for increasing equity, diversity, and inclusion in linguistics pedagogy. In *Proceedings of CLA 2020*, 1–13.
- Sandler, Wendy. 2012. The phonological organization of sign languages. *Language and linguistics compass* 6:162–182.
- Sandler, Wendy, and Diane C. Lillo-Martin. 2006. *Sign language and linguistic universals*. Cambridge University Press.
- Sehyr, Zed Sevcikova, Naomi Caselli, Ariel M Cohen-Goldberg, and Karen Emmorey. 2021. The ASL-LEX 2.0 project. *The Journal of Deaf Studies and Deaf Education* 26:263–277.
- Stokoe, William, Dorothy C. Casterline, and Carl G. Croneberg. 1965. *A dictionary of American Sign Language on linguistic principles*. Linstok.
- TED. 2014. TED^xGallaudet. <https://www.ted.com/tedx/events/9197>.
- Vicars, William G. 1997. ASL University. <http://www.lifeprint.com/index.htm>.
- Wood, David, Jerome S. Bruner, and Gail Ross. 1976. The role of tutoring in problem solving. *The Journal of Child Psychology and Psychiatry* 17:89–100. URL <http://dx.doi.org/10.1111/j.1469-7610.1976>.

tb00381.x.

Zuraw, Kie. 2022. Four inclusive practices for the phonology classroom.
In *Proceedings of the Annual Meetings on Phonology*, volume 9.

Reversing the gaze: Decolonizing the syllabus*

SOLVEIGA ARMOSKAITE
University of Rochester

Where do you begin telling someone their world is not the only one?
Lee Maracle, Stó:lō Nation¹

This paper reflects efforts of decolonization in higher education.² It offers a take on syllabus design by sharing some underlying principles and practices found effective over the years.

The paper draws on my experiences of learning and teaching. My graduate studies in the early 2000s at the Department of Linguistics at the University of British Columbia shaped me. There and then, Dr. Hotze Rullmann, the entire linguistics faculty, and the Indigenous language consultant Elders quietly, implicitly instilled in me principles that I have grown to embrace in my teaching practice. The paper makes the implicit practices explicit.

The practices reported here also rely on the never-ending dialogue that I have with my students at the University of Rochester (US). Striving to decolonize my own syllabi, I stumble across gaps in knowledge. For example, often an entire class does not know what boarding schools for Indigenous people were or has an entrenched idea that the local Indigenous population of Rochester or the entire New York state is long dead.

* Dr. Rullmann was on my dissertation committee and spent endless hours convincing me of possible worlds, semantically and metaphorically. I will always be grateful for his patience.

The title of this paper was provoked by a short speech by Dr. Kim TallBear opening a conference at the University of Alberta, where she stated that “It’s a long colonial tradition that Indigenous peoples are at the receiving end of the scientific gaze. It is time to reverse the gaze.” I quote from my notes; it comes from an eloquent and poignant YouTube video which has since become, unfortunately, unavailable.

¹ Lee Maracle (2017:61).

² This paper started as a well-received talk at the SUNY Council on Writing conference, fall 2021. The intense debate convinced me I should not have fretted that the content of the talk would be old news. As per their website, SUNY CoW is an institution with a long tradition, yet decolonization of the curricula is only being considered now: “Founded in 1980, the SUNY Council on Writing is an independent faculty organization dedicated to improving instruction and promoting scholarship in Writing and Rhetoric across the 64-campus system.”

The layout of the paper is as follows. We start with defining decolonization as part of transparent design (Section 1). Then we illustrate colonial bias and means to dislodge it through guided introspection and deploying tools provided by the Critical Language Awareness framework (Section 2). Next, we consider the significance of a specific geographical place as grounding a syllabus (Section 3). The discussion wraps up with appreciation of time and patterns of repetition (Section 4). Brief conclusions are offered last (Section 5).

1 Introduction: Why does a paper on pedagogy matter in the context of linguistics?

This section briefly explains what decolonization is understood to be (as there is more than one take on the matter) and why it is relevant to explicitly state the *modus operandi*.

1.1 What is decolonization of the syllabus?

A syllabus is part guide, part contract. It is partially a guide in that it lays out what will happen throughout the semester. It includes course policies, rules, and regulations, required texts and/or media, and a schedule as well as content of assignments. It can tell you nearly everything you need to know about how a course will be run and what will be expected. It is also partially a contract in that it recounts what the student and the instructor will deliver.

Yet syllabi transcend the student-instructor relationship. They also explicitly or implicitly (more on this in Section 2) manifest the culture and assumptions of the institution they represent. An overview of pedagogy literature³ and my own experience teaching diverse populations of students (such as first-generation college attendees, inmates at US high and medium security correctional institutions, international students, students from a range of privileged and/or underrepresented groups, etc.) converge on making *transparency* of syllabi and assignment design a priority. Here, transparency refers to explicit statements on course goals and learning objectives as well as the means deployed to achieve them.

For the narrow goals of this paper, transparency is discussed in relation to colonial versus de-colonized stance. That is, the question a reader of a syllabus should ask themselves is whether the syllabus

³ See Winkelmess et al. (2023) for a recent compilation.

reflects the values and needs of colonial curriculum — usually European white settler, often with Christian undertones — or whether it offers a means to serve and engage with the Indigenous, underrepresented, and marginalized.⁴ The (de)colonized stance can be implicit, whether intentional or not; it may be gleaned from conspicuous omissions or gaps, in, e.g., the reading list or particular obscure formulations in the assignment instructions. Said manifestation can also be explicit and intentional, revealing the rationale behind choices to hold up to the student the bias an instructor holds. Transparency in this matter doesn't just score higher or lower ratings on the scale of diversity and inclusion. It also humanizes the student-instructor relation by justifying, at the very least, the authority of the instructor as an expert while simultaneously offering an opportunity and a reason for critical engagement on the part of the student.

1.2 Why is it relevant to consider decolonization of syllabi in the context of linguistics?

While the decolonization of curriculum debate started decades ago,⁵ the efforts to implement the change are far from over (for great discussion and examples, see Figueroa 2020; Fuentes et al. 2021). In fact, many institutions and instructors are only beginning to catch onto the healing trend. Evidence of late or slow awakening abounds across institutional policies (e.g., only recently did we see a sweeping Indigenous land recognition momentum⁶) and practices (e.g., decolonizing talks and other

⁴ Granted, most syllabi fall on a continuum between the two extremes. This is due to the relative novelty of thinking about transparency of design in this capacity. Even the most well intentioned and driven instructors (current company included) are very much a product of decades of indoctrination by colonial curriculum; it will take time to successfully introspect and change our ways. Another reason for falling short may be lack of the necessary materials or knowledge about how to get such materials.

⁵ It depends on how or what one counts. If one goes by the oft-cited seminal work of Freire, *Pedagogy of the Oppressed*, the English translation first hit the shelves in 1970. If we consider Canada-specific Indigenous efforts, one could build on Battiste et al. 2002. Thinking globally, Māori scholar Smith's work on decolonizing methodologies is widely considered one of the foundational works, first published in 1999. The list could go on. What matters is that there is no longer a lack of relevant Indigenous materials that offer pedagogical frameworks and tools, yet these approaches have yet to infuse curricula of higher education across the fields.

⁶ Land recognition statements before events or on institutional policy pages are far from uncontroversial. The debate continues.

<https://www.npr.org/2023/03/15/1160204144/indigenous-land-acknowledgments>

materials are welcome and sought after⁷). Given that syllabi are manifestations of institutional policies and provide a modus operandi for courses, it is no wonder that forging syllabi more inclusive of the underrepresented is very much in vogue in the ivory towers.

Linguistics finds itself at the center of the decolonizing efforts due to what it is as an object of study (language intersecting with society) and as a tool for both communication (language) and meta-communication (language about language). Given that linguists have been involved in studying the languages of the colonized, we have over a couple hundred years of experience doing fieldwork which made us face the impact of colonization early on. Strides have been made to address the injustices brought to these communities by applying community-based participatory research methods (see Hacker 2013 for a classic overview of the framework); yet the work is far from over.

Here, the focus is narrowly on meta-communication about the content of the syllabi. Specifically, how one could/should/might use language when projecting or questioning a particular (de)colonized stance. While such considerations on meta-communication are necessary for designing linguistics syllabi, they are even more urgent in offering means for change in fields beyond linguistics.

2 Finding ways to engage

In this section, I give a sample of ways I use to dislodge the inertia of colonial thinking through some guided introspection and use of language.

2.1 Of pebbles, shells, and bookshelves

Given that the curriculum across disciplines remains largely Eurocentric, it also remains largely colonial. To address the disproportionate

⁷ Due to the page limit, I can only name but a few examples of Indigenous contributions for the sake of transparency. In the arts, the work of Canadians like Kent Monkman stands out (<https://www.youtube.com/watch?v=GwNpUevsKzc>). In literature, work by Thomas King (2017) rattles our conventions. The TED talk by Nigerian author Chimamanda Ngozi Adichie has gained a life of its own (<https://www.youtube.com/watch?v=D9Ihs241zeg>). In science, the hard-core activism of Kim TallBear (<https://kimtallbear.com/>) or the poetic ministrations of Robin Kimmerer (2013) come to mind. It is not possible to do justice to the wealth of what would be considered “hot” material available. And while this list of excellent Indigenous materials available is not exhaustive, looking for evidence of institutional change in higher education is exhausting, especially beyond the discipline of anthropology.

dominance, one could perhaps march into the classroom on day one and declare that this course will be decolonized henceforward. Declarations are hardly effective or convincing; if anything, one would run a risk of confusion (“Who, we? We are the good woke guys!”) or even resentment (“Oh yeah? I’m good, thanks. Make me!”). The means that I found to be effective are the opposite of public declaration. We start with an embodied silent introspection exercise, a version of an exercise on privilege that has been making the rounds online for a while now.⁸

The exercise goes as follows. At the onset of the class, students are asked to walk around the class where seven stations are set up. At each station, they silently read a vignette and have a choice to pick up either a pebble or a shell or both. Here are the vignettes:

- (1) Take a pebble if the language you speak at home is English. Take a shell if you speak an Indigenous language at home.
- (2) Take a pebble if you studied the history of the US from the settler perspective at school. Take a shell if you studied the history of the US from the perspective of Indigenous people.
- (3) Take a pebble if you can name at least three US writers or artists by name. Take a shell if you can identify at least three Indigenous writers or artists by name.
- (4) Take a pebble if you can name at least three US based inventions. Take a shell if you can name at least three Indigenous inventions.
- (5) Take a pebble if someone you look up to is a mainstream US icon. Take a shell if someone you look up to is an Indigenous icon.
- (6) Take a pebble if you can name three US holidays. Take a shell if you can name three Indigenous traditions.
- (7) Take a pebble if you can name three popular US dishes. Take a shell if you can name three Indigenous dishes.

⁸ I put together the content of the exercise (improvements and variations are endless). However, the exercise pattern itself was adapted from a workshop on integration of the LGBTQ community, and even there it has been adapted. I have searched the internet trying to get to the authorship of the exercise, but while versions and adaptations of the exercise abound, I have not been able to find the definitive source of origin.

At the end of the exercise, the students have a handful of items and, one hopes, a handful of thoughts. The first impact of the exercise lies in that it is embodied — one has to walk, touch, and collect. It would be strange to stand still and refuse to pick up an item, while it is really easy to ignore a declarative call to decolonize. The memory retains such a practice longer precisely because the whole body is involved. The contrast in touching and looking at pebbles and shells is deliberate, too: at some point in the discussion, I remind the participants that wampum shell beads were used as currency among some North American First Nations. The second impact of the exercise is that I never ask for what they have picked up or why. The recognition of gaps in knowledge is left private and lingers (or so the students tell me). We usually have a lively open-ended discussion when I ask them to share their thoughts after the exercise. This rudimentary check in on one's awareness of how colonial culture dominates and permeates our lives is effective as it reveals the lacuna of knowledge in a discreet private manner.

Another brief effective exercise that goes beyond individual education and reveals the bias in the setup of our knowledge systems is a library search for Indigenous material by Indigenous authors. The prompt is to ask students to find materials authored by, for example, Greek or Roman versus Indigenous authors using an online catalog of, for example, the local university library. The contrast is remarkable. The Eurocentric search gets hundreds of hits within seconds; yet it is nearly impossible to get a hit on an Indigenous author if one does not already know the name of the author and their tribal affiliation. The exercise in futility makes an impact on the discussion on what types of knowledge society legitimizes and prioritizes, often without making the choices transparent. By way of a wrap up, I share with the students my own efforts to get to Indigenous authorships and materials. It took me about four weeks and three librarians⁹ to get access to a couple of university library collections dedicated to Indigenous knowledge. This is all they found in North America. Here they are:

1. <https://guides.library.ubc.ca/aboriginalstudies/findingindigenouiperspectives>
2. <https://guides.library.queensu.ca/indigenous-studies/finding-authors>

⁹ Thanks to the University of Rochester librarians who found these collections: Stephanie Barrett, Eileen Daly-Boas, and Margaret Dull.

Both sites are hosted in Canada, and the University of British Columbia library takes the lead. That is, the Library of Congress is not changing their ways any time soon to accommodate alternative sources of knowledge.

2.2 Critical Language Awareness in the classroom

Once we have established, as discussed in Section 2.1, the awareness of how colonial our knowledge and curriculum still are, it is easier to transition into assignments for the course that transparently and deliberately deploy language to maintain the awareness and possibilities of other perspectives.

My assignment instructions and prompts are designed relying on the Critical Language Awareness (CLA) framework (I rely mainly on Shapiro 2022 and Curzan 2014), which provides linguistic tools to improve self-reflection, social justice, and rhetorical agency when creating or analyzing discourse patterns. I apply these tools to mindful use of language in the syllabi and assignments.

Due to constraints of space here and the wealth of literature that already exists, I would not be able to do justice to the CLA framework. I will rather walk the reader through two samples of how its tools can be deployed.

One way to deploy language with critical awareness is to consistently hold up and remind the students of the possible implicit bias or influencing factors that may be affecting their research. To that end, I include the following notice of consideration within my assignment prompts:

Consider if such factors as, e.g., your athletic ability, cultural or ethnic background, the education you received, gender, religious affiliation or social class etc. influence your experience and your research practice.

The hope is that the transparent direct request will make them introspect and that they will eventually pick up a habit of pondering the undercurrent of influences surfacing in their own thought process. I make sure to alert them if anything in their submitted work could be flagged as a factor.

In a similar fashion, dialogic thinking is encouraged through engagement with alternative views, often done in peer groups within in-class workshops. For example, if a student asks a research question such

as ‘*Why do Navajo retain their traditional sustainable practices even if it makes them fail in mainstream US economy?*’, an alternative question we forge in class might be ‘*Why does the settler economy fail to take into account the sustainable practices of the Navajo?*’ The hope is that the student sees how the tables can be turned depending on the worldview and that only by answering *both* questions can we get closer to the truth of the matter.

Another effective deployment of Critical Language Awareness is a comparative exercise that helps to tease apart patterns of thought and discourse in traditional Western scholarly texts and an Indigenous approach that may question the established patterns. A chapter from Kimmerer’s (2013) book begs for such a comparison. Kimmerer, as you may know, is a botanist who completed her Ph.D. in a mainstream North American university. Then she immersed herself in her native Potawatomi ethnobotany. The result is *Braiding Sweetgrass*, a book that fuses and compares both traditions, implicitly and explicitly.

The structure of the chosen chapter follows the traditional layout of a peer-reviewed research paper: it has an introduction, research question, methodology section, references etc. Yet the content of the chapter puts the Western approach to test. Due to constraints of space, we will look at two small excerpts from the chapter compared to two corresponding excerpts from an actual research paper.

Sample excerpt from the introduction to the chapter:

You can smell it before you see it, a sweet grass meadow on a summer day. The scent flickers on the breeze, you sniff like a dog on a scent, and then it’s gone, replaced by a boggy tang of wet ground. And then it’s back, the sweet vanilla fragrance, beckoning.

(Kimmerer 2013:156)

Sample excerpt from an introduction to a research paper:

At present only remnants of traditionally managed grassland biotopes, characterized by high biodiversity and a considerable conservation value (Kull & Zobel 1991; Garcia 1992) are found in Europe. Many characteristic species for these habitats are decreasing or threatened (Bastian & Bernhardt 1993).

(Losvik 2007:239)

Upon reading the two, students take a while to realize that the recounting of knowledge from sources in the piece of scholastic writing by the Swedish botanist Losvik is contrasted with sensory input from a naturalist in the milieu of the plant, Kimmerer, wearing her Potawatomi hat. That is, students are so conditioned to only expect Losvik-type style, that it takes them a moment to realize that the sensory statements of Kimmerer are noteworthy and built on generations of empirical observations. Usually, a great discussion ensues on whether the two modes of knowledge are comparable and complementary or not, and under which circumstances. Specifically, we consider if it is fair game to use sensory input such as the smell of approaching sweet grass as an introduction rather than providing a summary of scholastic sources on the subject matter.

The second set of short excerpts that leads to a marvelous discussion about clashing worldviews is the section of references from the two authors. Losvik lists her peer reviewed journal entries. Kimmerer lists sweet grass itself, her animal spirit, her collaborator student, and the ancestors.

References

Bastian, O. Bernhardt, A. 1993. Anthropogenic landscape changes in Central Europe and the role of bioindication. *Landscape Ecology*. 8: 139–151.

(Losvik 2007:247)

References

Wiingahsk, Buffalo, Lena, the Ancestors.

(Kimmerer 2013:166)

We ponder questions such as how and with what authority do we get to include sources of knowledge in the list of references? What assumptions do we have to share to justify the mention of sweet grass itself as a source of reference? Conversely, what assumptions do we have to make to exclude it?

My agenda is to show the contrast between the two worlds through choices in discourse and language. Whatever the students conclude for themselves, they at least have an encounter with an alternative worldview.

3 Exploring the place

Syllabi, as we have established, represent the institution. Institutions are largely outposts of colonizer knowledge, yet they are in geographical locations dotted with places of significance to Indigenous populations. Acknowledging these places and working their presence into the syllabi grounds the course materials and makes the Indigenous reality tangible for the students.

In the case of the University of Rochester, we address the fact that the campus we are on has been built on Haudenosaunee land. Depending on the course content and the objectives of the course, points of convergence between the course content and the recently built Seneca cultural center are found. If we are lucky enough to have money for the fieldtrip, we make the 20-minute journey to honor the Seneca heritage. For many, it is the first and only encounter with Indigenous culture in the context of their undergraduate studies.

If the course is on advertising, an introduction to Seneca culture contrasts with the chintzy imagery of the “noble Indian” that still permeates marketing ploys. The carefully curated and narrated exhibits contradict, for example, the repeated violations of Indigenous attire by brands like *Victoria’s Secret* or the appropriation of names by brands like *Jeep*.

If the course is on linguistics, we use an immersive experience into the intricacies of Seneca grammar through, for example, the ethnobotany trail constructed right outside the center. Here, a descriptive breakdown of plant names dovetails with detailed morphosemantic analysis of the words that identify the plants.

If the course is on writing, the possibilities are endless as the students come from different majors and chose to explore their own topics, ranging from engineers delving into the intricacies of constructing a Haudenosaunee longhouse to English majors attempting a comparative analysis of Western versus Seneca creation stories.

Repeated and consistent efforts to engage students with the local Indigenous heritage brings the historic and cultural footprint of the Indigenous community into academic discourse, to the here and now. It brings forward the vitality of Indigenous tradition.

4 Taking the time

In the previous section, I have illustrated the content as well as the means to ground the course materials within local Indigenous contexts. Last but

not least, time and timing are significant. An infusion of issues related to (de)colonization should seep through the topics and schedule outlined in a course syllabus. The assumption is that the majority of high school curricula encountered by students had either no or minimal authentic Indigenous thought. “*Repetitio est mater studiorum*” (“Repetition is the mother of learning”), says the Latin proverb. Therefore, we need to counter the repeated exposure to colonial curriculum with repeated consideration of Indigenous materials, or else we make no dent in the default colonial bias. Rather than devoting a single lecture to one specific topic, the effective way to absorb Indigenous thought would be to create a series of opportunities to re-examine the same materials throughout the course. In my practice, I provide at least three chances for engagement, which essentially results in a close reading of the text or deeper immersion in the media. In what follows, I walk the reader through one sample.

Take, for example, a writing course whose subject matter is Indigenous thought. Media and texts are compiled to introduce students to samples of Indigenous work in the arts and sciences. We start with a pre-writing module where we have a first take on the materials through classroom discussions guided by the students, which takes up about three classes. It is meant to be food for thought, a gentle encounter with the different worldview. A few weeks after the first encounter, we revisit the same compilation of materials with a different lens: we examine how Indigenous authors construct their arguments and provide evidence. This takes about two class periods. The third time we engage with the materials is to consider rhetorical choices in discourse. This takes about two class periods but might stretch out depending on individual needs, as individual instructor-student conferences occur at this time in the course. Thus, the same materials are scrutinized at least thrice, with a deepening understanding and attention to detail, deploying a variety of analytical tools. The result is intimate knowledge of a sliver of Indigenous thought. The hope is that the planted seed takes root, and the students retain at least some curiosity to explore further.

5 The syllabus is the locus of a clash in worldviews

We argued that the syllabus is where the dismantling of the pervasive colonial legacy is set off. We started with introspection of our own bias followed by inspection of the bias evident in the legitimized systems of knowledge, such as the organizing principles of libraries. Next, we illustrated how Indigenous thought can be effectively contrasted with

colonial inertia. Specifically, we advocated for transparent design. Combined with tools from the Critical Language Awareness framework, transparent design helps us craft syllabi and assignments with explicit intent and poignant discourse choices. Imbuing syllabi with locally significant geographical and cultural features makes Indigenous heritage come alive to the students. Giving students the time to re-examine materials through shifting lenses across several weeks ensures intimate knowledge. Thus, we showed that in an academic environment, the syllabus is “where you begin telling someone their world is not the only one”, to steal the line from Lee Maracle.

References

- Battiste, Marie, Lynne Bell, & L. M. Findlay. 2002. Decolonizing education in Canadian universities: An interdisciplinary, international, indigenous research project. *Canadian Journal of Native Education* 26(2):82–95.
- Curzan, Anne. 2014. *Fixing English: Prescriptivism and language history*. Cambridge, UK: Cambridge University Press.
- Figueroa, Michael A. 2020. Decolonizing “Intro to World Music?”. *Journal of Music History Pedagogy* 10(1):39–57.
- Freire, Paulo. 1970 [2000]. *Pedagogy of the oppressed*. New York City, NY: Continuum.
- Fuentes, Milton A., Zelaya, David G., & Joshua W. Madsen. 2021. Rethinking the course syllabus: Considerations for promoting equity, diversity, and inclusion. *Teaching of Psychology* 48(1), 69–79.
- Hacker, Karen. 2013. *Community-based participatory research*. Thousand Oaks, CA: Sage Publications.
- Kimmerer, Robin. 2013. *Braiding sweetgrass: Indigenous wisdom, scientific knowledge and the teachings of plants*. Minneapolis, MN: Milkweed editions.
- King, Thomas. 2017. *The inconvenient Indian illustrated: A curious account of native people in North America*. Toronto, ON: Doubleday Canada.
- Lee, Maracle. 2017. *Ravensong*. Toronto, ON: Canadian Scholars’ Press.

- Losvik, Mary H. 2007. Regional species pools of hay meadows: a case study. *Applied Vegetation Science* 10(2):239–248.
- Shapiro, Shawna. 2022. *Cultivating critical language awareness in the writing classroom*. London, UK: Routledge.
- Smith, Linda T. 2021. *Decolonizing methodologies: Research and indigenous peoples*. New York, NY: Bloomsbury Publishing.
- Winkelmes, Mary-Ann, Allison Boye, & Suzanne Tapp (eds.). 2023. *Transparent design in higher education teaching and leadership: A guide to implementing the transparency framework institution-wide to improve learning and retention*. New York, NY: Taylor & Francis.

Media

<https://ganondagan.org/>

<https://www.npr.org/2023/03/15/1160204144/indigenous-land-acknowledgments>

<https://www.youtube.com/watch?v=GwNpUevsKzc>

<https://www.youtube.com/watch?v=D9Ihs241zeg>

<https://kimtallbear.com/>

Celebrating and quantifying the linguistic diversity of the UBC student community

MOLLY BABEL, LINE LLOY, ZARA KHALAJI PIRBALUTI,
RACHEL SOO & LEXIA SUITE
University of British Columbia

1 Introduction

While the University of British Columbia (UBC) does not survey the student body's language background, the undergraduate student body is a diverse lot.¹ The goal of this short paper is to provide a first-pass description of the linguistic diversity in the student population at UBC.

That UBC would be linguistically diverse is unsurprising given the surrounding speech community. Metro Vancouver boasts high levels of linguistic diversity itself. For example, while English is the dominant societal language of both the university and Vancouver, only 51.2% of Metro Vancouver residents are mother tongue speakers of English (ISO 639-3: eng). French (ISO 639-3: fra) is not widely spoken as a mother tongue in Metro Vancouver, with less than 2% speaking French as their mother tongue. Over fifty different mother tongue languages are spoken by the non-English and non-French mother tongue speakers, according to the most recent census (Statistics Canada, 2023).

In our characterization of the linguistic diversity of UBC students, we consider various aspects of the language experience and quantify language patterns from several angles. This multi-pronged approach is in recognition that bilingualism — or, more broadly, multilingualism — is a challenging, if not impossible, construct to quantify (Marian and Hayakawa, 2021). Moreover, we highlight that any measure of bilingualism is a continuum and not a categorical variable (Luk and Bialystok, 2013). A common instrument used to describe individuals' multilingual experiences is the Language Experience and Proficiency Questionnaire (LEAP-Q, Marian et al., 2007). The LEAP-Q probes participants' lan-

¹ Thank you to the UBC community for sharing your language background with us. Thanks to Khia A. Johnson and Khushi Nilesh Patil for their contributions to the projects from which these data originate. We thank Hotze Rullmann for being such a wonderful teacher and colleague! We are lucky to have you in our lives.

guage history, use, attitudes, and self-rated proficiency, providing data that can be quantitatively or qualitatively described.

Using responses on the LEAP-Q, Gullifer and Titone (2020) recently introduced a measure called *language entropy* that quantifies the predictability of an individual's language use in different contexts. In the quantification of language entropy, a monolingual individual would have a score of 0 in any context; there is no doubt about the language that will be used, as the individual is monolingual. A bilingual individual who uses both of her languages equally in a given environment would have an entropy value of 1, indicating that it is unpredictable which of the two languages would be used. The maximum entropy value increases with the number of languages spoken, but, regardless of the number of languages spoken, a low language entropy value indicates that it is highly predictable what language that individual would use in a given context and a high language entropy value indicates unpredictability in language use. Gullifer and Titone (2020) characterize these types of language use associated with low and high language entropy as *compartmentalized* and *integrated*, respectively, pointing to the ways in which an individuals' multiple languages are used in varying social contexts. As a kind of validation of this interpretation, language entropy is positively correlated with language mixing and switching practices (Kałamała et al., 2022), though it appears to be independent from cognitive processing measures like proactive control (Wagner et al., 2023; Gullifer and Titone, 2021).

The goal of this paper is to provide a description of the multilingualism of UBC students. Because we intend for this paper to be broadly readable, we avoid quantitative analyses and, instead, provide qualitative descriptions of the patterns.

2 Methodology

2.1 Participants

1026 UBC students completed the LEAP-Q. Ten individuals did not report their month and year of birth. The mean participant age was 22 (SD = 3.7). As this is a rather contracted age range, we do not discuss age further.²

² We note changes in language use over time may be an interesting and meaningful dimension to consider, should the data allow.

2.2 Materials and procedures

The LEAP-Q was administered on Qualtrics. For the subset of data from Suite et al. (2023), this instrument was presented after a short vocabulary assessment in a survey that followed completion of a sentence transcription task. For the subset from Lloy et al. (2024), the LEAP-Q was completed in a multilingual survey that also included the Bilingual Language Profile (Gertken et al., 2014) and the Bilingual Code Switching Questionnaire (Rodriguez-Fornells et al., 2012). In both projects, the LEAP-Q was completed by participants online in a location of their choosing.

3 Results

3.1 What type of multilingual?

Figure 1 presents two panels that broadly summarize the type of multilingual speakers in the UBC speech community. On the left, Panel A is a histogram of the number of individuals who report experience with different numbers of languages. The mode of this distribution is 3, indicating the most common situation is to have experience with three languages. Bilingual and quadrilingual experiences are the next most likely language backgrounds. It is more common for UBC students to have experience with five languages than to be monolingual.

An important distinction in the bilingual (or multilingual) experience is whether an individual acquired their first two languages simultaneously or sequentially. Sequential bilinguals who learn a second language much after their first often, but not always, exhibit linguistic patterns distinct from simultaneous bilinguals. To determine whether UBC students are simultaneous or sequential bilinguals, the two lowest reported ages of acquisition for individuals with experience with more than two languages were compared. The difference in these values is reported on the x-axis in the right panel of Figure 1. There is a large spike at 0, indicating that the mode is for individuals to be simultaneous bilinguals; there is no difference in the ages at which individuals begin acquiring their first two languages. A second clear peak in the data occurs before the onset of schooling. As most participants report age 0 as the onset of acquisition of their first language, this second peak in early childhood may suggest that many individuals begin acquiring a second language in an early childcare

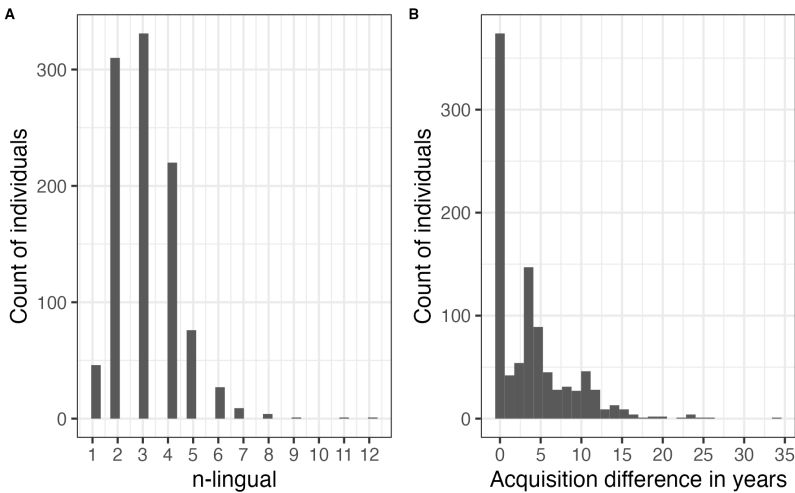


Figure 1: (A): A histogram of how multilingual participants are. The vertical axis shows the counts of individuals for each n -lingual bin on the horizontal axis. Trilinguals are the most common type of multilingual. (B): A histogram of the difference in the ages at which individuals acquire their first two languages. An acquisition difference of 0 represents simultaneous bilinguals, for whom there is no difference in age of acquisition of their first two languages.

or educational setting either due to entrance in a language immersion program or an introduction to English in daycare or preschool. English, the societally dominant language in the Lower Mainland, is then introduced at this point after having familial experience with another language.

Our calculation of language entropy provides separate values for speaking/signing³, exposure, and reading, as individuals can vary in how often they produce a language, how often they are exposed to language, and how often they read a language. These varied experiences are observed in the panels in Figure 2, which shows speaking by exposure entropy, and Figure 3, which shows speaking and exposure entropy by reading entropy. Because of an interest in characterizing different calculations of entropy, particularly speaking and exposure entropy, we present these data in scatterplots that show pairwise correlations and histograms along the top and right sides of the figures.

³ We use the term ‘speaking’ for any kind of oral or signed language production.

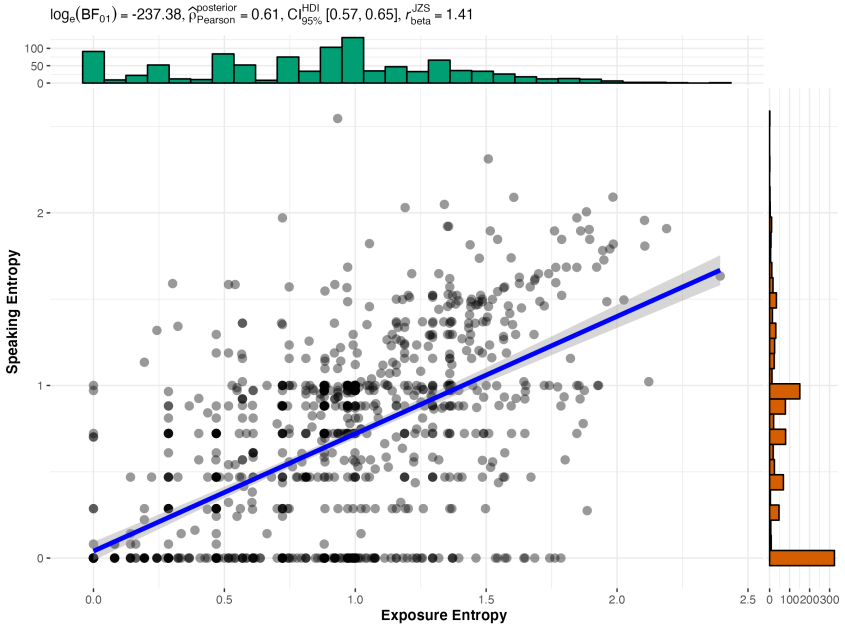


Figure 2: Speaking Entropy (vertical axis) by Exposure Entropy (horizontal axis). Histograms for both variables are on the opposing axes.

Each pairwise comparison demonstrates a positive correlation. This suggests that individuals with high entropy for speaking/signing, exposure, or reading are also more likely to have high entropy values for any of these dimensions. So, while we see from the histograms in these figures that, for example, there is a more prominent low entropy peak for reading and speaking than exposure, the overall pattern in these values is that more integrated language use in one domain is associated with more integrated language use in another domain. However, the strength of the relationship is the strongest for speaking and exposure entropy, suggesting that reading is a more distinct mode.

3.2 What languages are represented?

Having established that UBC students have experience with multiple languages, let us identify what those languages are.

Participants reported speaking 104 distinct languages. This language

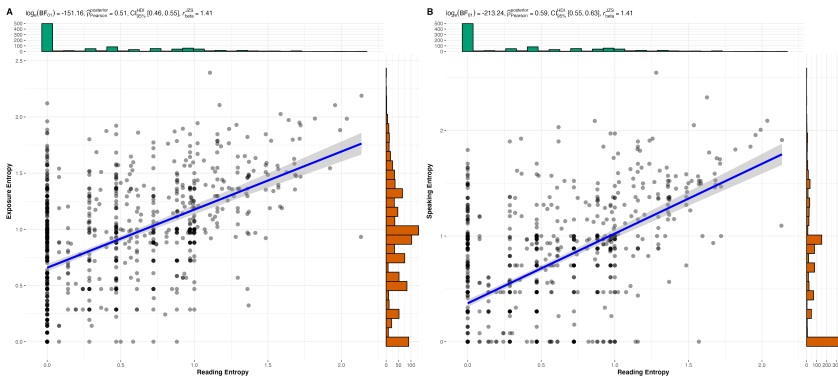


Figure 3: (A) Exposure Entropy (vertical axis) by Reading Entropy (horizontal axis). (B) Speaking Entropy (vertical axis) by Reading Entropy (horizontal axis). In both panels, the histograms for both variables are on the opposing axes.

diversity was attenuated in participants' reports of their most dominant language; there were 24 languages reported as participants' most dominant. The linguistic diversity increased for individuals' second (59 reported languages) and third most dominant languages (54 reported languages). The 15 most commonly spoken languages in each of these groups — all languages, and the most, second, and third dominant languages — are reported in Table 1. English dominates the column for all reported languages and dominant languages. This is unsurprising given that the language of instruction at UBC is generally English. French is the most often reported second and third most dominant language, but only 4 individuals report French as their most dominant language. This presumably is a UBC manifestation of the mother tongue census data in BC, which reports that less than 2% of BC residents are mother tongue speakers of French.

With the exception of English, the number of speakers reporting a particular language as a second most dominant language outnumber the count of individuals who report that same language as a most dominant language. This asymmetry is likely indicative of the rich diversity in home languages in our domestic student population. The home language environment is ultimately usurped by English dominance due to the soci-

etal dominance of English.

The languages reported for the top three dominant languages and all languages are presented visually in word clouds in Figure 4. Interactive versions of these four figures are available for download [here](#).⁴ English has been removed from these visualizations since its size inhibits the readability of the other languages.

Meriting special mention are the four First Nations languages reported in our sample: Anishinaabemowin (ISO 639-3: oji/ojg), Chinuk Wawa (IISO 639-3: chn), hən̓q̓əmiñə́m̓ (ISO 639-3: hur), and Nehiyawewin (ISO 639-3: crk). We celebrate their presence and hope to see an increase in the number of First Nations languages spoken by our student body in future years.

Table 1: The 15 most frequently reported languages and their reported counts, ordered by frequency. The columns present the languages most commonly reported overall (first column), and those most commonly identified as the dominant language (second column), as the second most dominant language (third column), and as the third most dominant language (fourth column).

All languages	Dominant Language	2nd Dominant Lg	3rd Dominant Lg
English (1025)	English (841)	French (190)	French (210)
French (525)	Mandarin (73)	English (165)	Mandarin (94)
Mandarin (348)	Cantonese (36)	Mandarin (131)	Spanish (78)
Spanish (210)	Korean (23)	Cantonese (103)	Japanese (50)
Cantonese (184)	Japanese (8)	Spanish (59)	Cantonese (31)
Japanese (151)	Russian (6)	Korean (41)	Hindi (31)
Korean (129)	French (4)	Punjabi (38)	English (19)
Hindi (66)	Hindi (4)	Tagalog (34)	Korean (19)
Punjabi (57)	Spanish (4)	Japanese (31)	German (17)
German (54)	Farsi (3)	Hindi (23)	ASL (13)
Tagalog (51)	Punjabi (3)	Arabic (20)	Italian (12)
Arabic (38)	Tagalog (3)	German (15)	Tagalog (10)
Russian (30)	Turkish (3)	Portuguese (12)	Punjabi (9)
Italian (26)	Arabic (2)	Vietnamese (11)	Hokkien (8)
ASL (23)	Bahasa Indonesian (2)	Farsi (9)	Russian (6)

⁴ On the interactive html files, moving one’s cursor over the language name shows the number of individuals who reported that language.

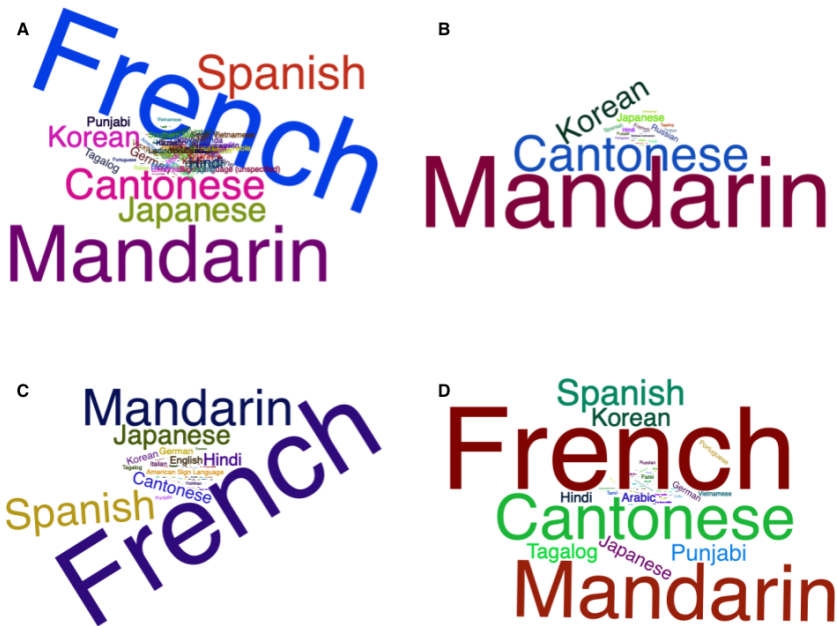


Figure 4: Language frequency word clouds. Word clouds, from right to left, top to bottom, visualizing the frequency of (A) all languages in the data set, (B) most dominant languages, (C) second most dominant languages, (D) third most dominant languages. English is excluded from all clouds.

3.3 Self-ratings

Individuals provided self-proficiency ratings for speaking and understanding in each of their languages. These data are provided in Figure 5 for up to the six most dominant languages. Self-ratings for speaking and understanding are at ceiling for Language 1 (individuals' most dominant language) and gradually lower as the language becomes less dominant.⁵ The second and third most dominant languages demonstrate an interest-

⁵ The individual data points presented as circles in these boxplots represent responses that are aberrant with respect to the general response distribution. In some cases, these data points are due to a likely misreading of the survey; individuals were asked to enter in their languages in the order of dominance, but some entered their languages in the order of acquisition.

ing asymmetry in language use: Individuals report higher proficiency in understanding than speaking. This difference does not exist for the most dominant languages or fourth most dominant languages and above.

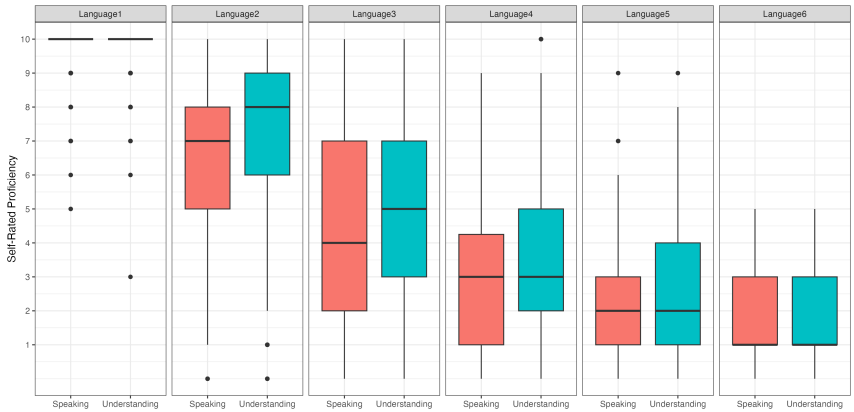


Figure 5: Boxplot visualization of the self-ratings for individuals' top six most dominant languages.

4 Conclusion

This short paper represents a first attempt at describing the linguistic diversity in the student population at UBC. Using questionnaire data collected from over 1000 UBC students (Suite et al., 2023; Lloy et al., 2024), we provide a qualitative description of the types of multilinguals, their various language entropy scores, and the languages they speak.

While we should celebrate the linguistic diversity of our UBC students, this multilingual profile is not unique. The majority population in the world is multilingual (Grosjean, 2021). At the same time, monolingual speakers are often placed on a pedestal, as though their linguistic competence and performance is more authentic than that of a multilingual speaker (Cheng et al., 2021). In celebrating the linguistic variation of UBC students, we also showcase the opportunity to innovate discipline-moving research questions that improve our theory and understanding of linguistic knowledge, behaviour, and processes.

References

- Cheng, L. S., Burgess, D., Vernooij, N., Solís-Barroso, C., McDermott, A., and Namboodiripad, S. (2021). The problematic concept of native speaker in psycholinguistics: Replacing vague and harmful terminology with inclusive and accurate measures. *Frontiers in Psychology*, 12:715843.
- Gertken, L. M., Amengual, M., and Birdsong, D. (2014). Assessing language dominance with the Bilingual Language Profile. In Leclercq, P., Edmonds, A., and Hilton, H., editors, *Measuring L2 proficiency: Perspectives from SLA*, pages 208–225. Multilingual Matters, Bristol, UK.
- Grosjean, F. (2021). *Life as a bilingual: Knowing and using two or more languages*. Cambridge University Press, Cambridge.
- Gullifer, J. W. and Titone, D. (2020). Characterizing the social diversity of bilingualism using language entropy. *Bilingualism: Language and Cognition*, 23(2):283–294.
- Gullifer, J. W. and Titone, D. (2021). Engaging proactive control: Influences of diverse language experiences using insights from machine learning. *Journal of Experimental Psychology: General*, 150(3):414.
- Kałamała, P., Senderecka, M., and Wodniecka, Z. (2022). On the multidimensionality of bilingualism and the unique role of language use. *Bilingualism: Language and Cognition*, 25(3):471–483.
- Lloy, L., Patil, K. N., Johnson, K. A., and Babel, M. (2024). Language-general versus language-specific processes in bilingual voice learning. *under review*, pages 1–45.
- Luk, G. and Bialystok, E. (2013). Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *Journal of Cognitive Psychology*, 25(5):605–621.
- Marian, V., Blumenfeld, H. K., and Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, 50(4):940–967.

- Marian, V. and Hayakawa, S. (2021). Measuring bilingualism: The quest for a “bilingualism quotient”. *Applied Psycholinguistics*, 42(2):527–548.
- Rodriguez-Fornells, A., Krämer, U. M., Lorenzo-Seva, U., Festman, J., and Münte, T. F. (2012). Self-assessment of individual differences in language switching. *Frontiers in Psychology*, 2:388.
- Statistics Canada (2023). (table) Census Profile. 2021 Census of Population. Statistics Canada Catalogue no. 98-316-X2021001. Ottawa.
- Suite, L., Freiwirth, G., and Babel, M. (2023). Receptive vocabulary predicts multilinguals’ recognition skills in adverse listening conditions. *The Journal of the Acoustical Society of America*, 154(6):3916–3930.
- Wagner, D., Bekas, K., and Bialystok, E. (2023). Does language entropy shape cognitive performance? A tale of two cities. *Bilingualism: Language and Cognition*, pages 1–11.

Mood in Wá·šiw narratives: A first look*

M. RYAN BOCHNAK

University of British Columbia

1 Introduction

This paper presents a first look at the use of mood morphemes in narratives in Wá·šiw (also known as Washo, Washoe; isolate; California and Nevada, USA). I examine the use of mood morphemes in four versions of the story *Coyote and Lizard* as told by four different Wá·šiw speakers across time, with special attention given to the independent mood *-i* and dependent mood *-aʔ*. I compare the distribution of these mood markers found in the texts with the predictions of the analysis of mood markers by Bochnak and Hanink (2022) and Bochnak (2023). Of particular interest is the distribution of the dependent mood marker *-aʔ*, which is referred to by Jacobsen (1964) as a “narrative tense”, a label suggestive of its relative prominence in narratives and texts. And indeed, this study confirms that the distribution of *-aʔ* in narratives is wider than predicted under recent theoretical accounts. Some avenues for possible ways of analyzing the narrative use of the dependent mood are explored, and I suggest that a comparison with the reportative subjunctive in German may be apt.

2 The mood system in Wá·šiw and predictions of previous analyses

In Wá·šiw, there is a set of inflectional morphemes that appear towards the end of the verbal complex and form a finite clause. These are called “final suffixes” by Jacobsen (1964), since only nominalizing or adverbializing morphology and switch reference can appear to their right. These morphemes are re-cast as moods by Bochnak (2016), and this label con-

* Congratulations on your birthday Hotze! *gahamuʔaŋawšému!* Thank you for being a wonderful colleague and mentor. The topic for this paper was inspired by Hotze’s interest in the grammar of narratives, and by Reisinger, Matthewson, and Rullmann’s (2022) corpus study of modals to test the predictions made by Rullmann and Matthewson (2018). Funding from a Hampton Research Grant and a Mellon Foundation Fellowship is gratefully acknowledged.

tinues to be used in subsequent work.¹ The set of mood morphemes is shown in Table 1.

Table 1: Moods in Wá·šiw

<i>-i</i> independent	<i>-hi</i> optative	<i>-le</i> redundant
<i>-aʔ</i> dependent	<i>-hulew</i> hortative	$-\emptyset$ imperative

At issue in this paper is the distribution of the independent mood marker *-i* and the dependent mood marker *-aʔ*. In contrast to the other moods, which are used for expressing certain semantic notions related to the modal and/or informational status of a clause (Bochnak 2023), the distribution of the independent and dependent moods is largely predictable based on the syntactic environments in which they appear. According to Bochnak and Hanink (2022), the dependent mood *-aʔ* is used in various types of adjunct clauses, such as temporal adjunct clauses with a temporal overlap interpretation, concessive clauses which convey some sort of contrast with the matrix clause, and complement clauses in non-factive attitude reports. Meanwhile, the independent mood *-i* is the default mood for matrix clauses, and is required in certain subordinate clauses, such as relative clauses and complement clauses in factive attitude reports. This distribution is summarized in Table 2.

Table 2: Distribution of *-i* and *-aʔ* (adapted from Bochnak 2023)

	<i>-i</i>	<i>-aʔ</i>
matrix clauses	✓	*
relative clauses	✓	*
factive complements	✓	*
non-factive complements	*	✓
temporal adjunct clauses	*	✓
concessive clauses	*	✓

In Bochnak and Hanink's analysis, they treat the independent mood as the elsewhere case, inserted whenever another mood marker, including the dependent mood, cannot be. In particular, the default mood for matrix clauses is the independent, whereas the dependent is restricted to

¹ See Bochnak (2023) for explicit argumentation that these should be treated as moods in the sense of Portner (2018).

subordinate clauses. Problematic for this analysis is the qualitative observation that the dependent mood appears to have a wider distribution in narratives, where it appears, at least on the basis of translations, to be used in matrix clauses.² As Jacobsen writes: “it [the dependent mood — MRB] is often used as a narrative tense, the tense in which tales are told” (1964: 663). Jacobsen’s comments suggest that the dependent mood may even be the default for narratives.

In this paper, I take the first steps towards investigating the use of the independent and dependent moods in Wá·šiw narratives. Specifically, I quantify the use of these moods in four versions of the *Coyote and Lizard* story. My goal is to uncover to what extent the distribution of these moods conforms to Table 2, and to what extent the distribution of the dependent mood *-a?* occurs outside of those environments, specifically in (apparent) matrix clauses.

3 The texts and methodology

For this study, I examined four versions of *Coyote and Lizard*, which is a folklore tale that explains why human hands are shaped the way they are. Coyote and Lizard argue with each other over what type of hands humans should have. After some shenanigans, Lizard is ultimately the winner, and humans have hands that look similar to Lizard’s with extended fingers rather than Coyote’s paws.

In Table 3, I give the following information for each text: the speaker who told the story orally, the date it was recorded, the person who collected the story, and the total number of clauses in the narrative. I also include a reference code that is used in example sentences to refer to the version of the text it comes from. Example sentences from texts also contain the sentence number that they come from.

This particular set of texts was chosen for many reasons. First, they are versions of the same story, which mitigates possible effects of genre as an influence of mood choice. Second, there is a balance between male and female speakers. Third, there are two generations of speakers represented, with two texts being recorded in the 1950s, and two being recorded in the late 1990s to early 2000s. Fourth, these texts all already had morphological parses available, though the level of detail of those parses varied.³

² As such, this use of the dependent mood, which is otherwise restricted to subordinate clause types, appears to be a case of insubordination (Evans 2007).

³ Thanks to Emma Wilcox and Alan Yu who completed some of the parses with

Table 3: Versions of *Coyote and Lizard* used in this study

Code	Speaker	Date	Collected by	Total # of clauses
JW	John Wiger	29 Nov 1955	William Jacobsen	142
BH	Bertha Holbrook	18 July 1956	William Jacobsen	112
SA	Sylvia Andrews	ca. 1997	Laura Fillmore	50
SJ	Steven James	20 Aug 2004	Alan Yu	42

For each text, I counted the total number of matrix and subordinate clauses, the total number of clauses containing the independent *-i* and the dependent *-aʔ*, the number of matrix and subordinate clauses containing moods *-i* and *-aʔ*, and whether each use of *-i* or *-aʔ* is predicted by the distribution given in Table 2. Arriving at the final numbers was not an easy task or an exact science. Namely, the classification of matrix vs. subordinate clause is already an analytical choice. Since the research question of this paper asks to what extent the dependent mood *-aʔ* appears outside of its predicted distribution, I could not use the presence of *-aʔ* as indicating that a clause is necessarily subordinate. Instead, I considered a number of factors — both in the morphology of the clauses and their interpretation — to arrive at the count of matrix and subordinate clauses with each mood type. Those factors essentially align with the summary in Table 2 of the distribution of the independent and dependent moods.

For independent *-i*, its presence was predicted in the context of relative clause marking (*-gi* or *-ge*), sequential marking (*-ud*), or other overt markers of subordination (e.g., the adverbializer *-da*). These are cases where *-i* occurs in subordinate clauses,⁴ and where the dependent *-aʔ* is not possible. Beyond that, the independent mood is predicted in matrix clauses. The sentence in (1) shows both an instance of *-i* in a subordinate clause and an instance of *-i* in a matrix clause. The first *-i* in the sentence appears in a clause that also contains relative clause marking (*-gi*). This form of subordination is used in (1) to mark the complement of the modal verb *-eʔ*. The independent mood *-i* is correctly predicted, since it is the only mood marker that can co-occur with *-gi* (Jacobsen 1964). The matrix clause also shows the default use of *-i* in matrix clauses, in this case

me.

⁴ The term ‘independent’ simply refers to the fact that this mood is the default in matrix clauses, even though this mood can also appear in certain types of embedded clauses, as shown in Table 2.

marking the modal *-eʔ*.⁵

- (1) Context: Lizard says that humans will be like him and have five fingers. Coyote disagrees and says,

lé:duŋ	ʔéʔuʃgabigi	k'éʔi
le:-duŋ	ʔ-eʔ-uš-gab-i-gi	k'-eʔ-i
1-like	3-be-DUR-DIST.FUT-IND-SUBJ.REL	3-MOD-IND

'They will be like me.' (BH: 8)

For dependent *-aʔ*, the situation is a bit trickier, since this mood is largely used by itself to mark that a clause is subordinate, rather than co-occurring with overt subordinating morphology. Thus, there are many clauses containing *-aʔ* where there is no other morphological marking that unambiguously identifies a clause as subordinate. An instance of *-aʔ* was counted as being predicted by Table 2 under the following conditions. First, if it occurred in a clause to the left of a clause whose main verb was a verb of thinking or saying, indicating that the *-aʔ*-marked clause is a non-factive complement. Such an example is given in (2), where the clause marked with *-aʔ* is the complement of *ʔi:demelʔgi* 'they said'.⁶

⁵ The orthography used is slightly modified from Jacobsen (1964), where most characters have their typical IPA value, with the following exceptions: M = [m̥], š = [ʃ], y = [j]. The colon : represents a long vowel. I use the following glosses: 1, 3 = 1st, 3rd person; CAUS = causative; DEP = dependent mood; DIST.FUT = distant future; DIST.PAST = distant past; DUR = durative; IND = independent mood; INS = instrumental; LOC = locative; MOD = modal; PAST = past tense; PRO = pronoun; Q = question; RED = reduplication; REDUND = redundant mood; REFL = reflexive; SEQ = sequential; SR = switch reference; SUBJ.REL = subject relative; THEME = anaphoric theme; TOP = topic change; TRAD = traditional.

⁶ The use of the subject relative marker *-gi* at the end of the final clause in (2) also appears to be a case of insubordination, i.e., a morphologically subordinate form used as a main clause. The clause in which it appears does not seem to be subordinate to any clause within (2) or in the next sentence that follows in the text.

- (2) Context: First sentence of the text

wi:diʔ pítelihak'a géwe
 wi:diʔ piteliʔ-hak'a gewe
 this lizard-with coyote

guMitgá:k'ululiyaʔ
 Ø-guM-itga:k'u-lul-li-aʔ
 3-REFL-disagree-DIST.PAST-long.ago-DEP

ʔi:demelʔgi
 ʔ-i:d-emelʔ-i-gi
 3-say-TRAD-IND-SUBJ.REL

'They said that this lizard and the coyote had a disagreement with each other.' (JW: 1)

Second, to identify temporal adjunct or concessive uses of *-aʔ*-marked clauses, I looked for English translations that included “while”, “when”, “as”, or “but”, or a gerund verb form indicating temporal simultaneity with a superordinate clause. (3) represents such an example, where the first clause marked with *-aʔ* is interpreted as a temporal adjunct clause.⁷ Meanwhile, the second clause was counted as a case of a matrix use of *-aʔ*, since it does not appear to be in an obvious subordinate relationship with any other clause in this sentence or in the following sentence in the text.

- (3) Context: Lizard and Coyote are arguing.

píteliʔ Múʔšamušgap'ilaš géwe
 piteliʔ Ø-Mu-iʔiš-am-uš-gap'íl-aʔ-š gewe
 lizard 3-run-forward-away-DUR-here.and.there-DEP-SR coyote

galóʔpamduwéweʔaʔ
 ge-loʔop-am-duweweʔ-aʔ
 3.OBJ-catch-away-try.RED-DEP

'While Lizard ran away, Coyote chased after him.' (SJ: 5)

I also counted an *-aʔ*-marked clause as a temporal adjunct clause if the action or state named in the clause was plausibly occurring simultaneously with the action or state named in a neighbouring clause, even if the translation did not contain one of those aforementioned lexical items.

⁷ As explained below, the presence of switch reference *-š* in this clause also indicates that it is a subordinate clause.

This means that at least some cases were somewhat up to interpretation.⁸ In (4), I counted the first clause marked with *-a?* as a temporal adjunct clause: Coyote being pleased with himself very plausibly occurs simultaneous to his thinking that he has burned Lizard (the latter being the likely reason that he is pleased with himself).

- (4) Context: Lizard runs under a rock to hide from Coyote. Coyote stuffs sagebrush under the rock and sets it on fire in order to burn Lizard.

ʔudi	géwe	gumgaʔlá:maʔ	gik
ʔ-ud-i	gewe	∅-gum-gaʔla:m-aʔ	gik
THEME-SEQ-IND	coyote	3-REFL-like-DEP	3.PRO
dót'ikhayaʔ	hámuyaʔ	ʔi:yeweʔi	
∅-dot'ig-ha-aʔ	∅-hamu-aʔ	ʔ-i:yeʔ-uweʔ-i	
3-burn-CAUS-DEP	3-think-DEP	3-go-hence-IND	

‘Then Coyote left, thinking that he had burnt [Lizard] and was pleased with himself.’ (JW: 17)

In fact, I argue that every mood marker in (4) is predicted by the generalizations in Table 2. In addition to the dependent mood marking on *gumgaʔlá:maʔ* ‘he was pleased with himself’ as indicating a temporal adjunct clause, I argue that the *-a?* marking on the *hámuyaʔ* ‘he thought’ is also a temporal adjunct use of the dependent mood: Coyote being pleased with himself and his thinking occur simultaneously with the action of the matrix clause, his going away. The verb form *dót'ikhayaʔ* containing *-a?* ‘he burnt him’ is a complement clause of the verb *hámuyaʔ* ‘he thought’ (i.e., ‘he thought that he had burnt him’). Finally, the independent mood *-i* on the final clause is the default for matrix clauses.

An *-a?*-marked clause containing overt switch reference morphology also counted as a subordinate clause, since switch reference marking in Wá·šiw is a property of subordinate clauses only (Arregi and Hanink 2022). Switch reference marks disjoint reference of the subjects in the matrix and subordinate clauses. (5) is such an example.⁹ The first clause is marked with switch reference since the third person subject of the first clause (Coyote) is disjoint with the third person subject of the second clause (Lizard). In this particular example, it is plausible that the first

⁸ Since I tagged the clauses myself, future studies would ideally have at least one more annotator so that inter-annotator agreement could be checked.

⁹ Switch reference marking can also occur on subordinate clauses marked with independent *-i*.

clause is interpreted as a temporal adjunct clause (i.e., “When Coyote walked away from there, Lizard came out from there.”). In any case, the switch reference marking on the first clause unambiguously signals that this is a subordinate clause. I also counted the second clause in this sentence as a temporal adjunct clause, because the event it described plausibly occurs simultaneously with the event of the following clause, namely that Lizard taunts Coyote.

- (5) Context: Lizard crawls under a rock, and Coyote tries to burn him. After a while, Coyote gives up and walks away.

yá: p’íʔšuweʔuŋilaš píteliʔ dášiʔ
ya: Ø-p’-iʔiš-uweʔ-uŋil-aʔ-š piteliʔ da:-šiʔ
then 3-crawl-forward-hence-PAST-DEP-SR lizard there-from
pímiʔaʔ
Ø-p’-imiʔ-aʔ
3-crawl-out.from-DEP

‘He [Coyote] walked away from there. Lizard came out from there [under the rock].’ (BH: 26–27)

I counted any other *-aʔ*-marked clause found in a text as a matrix use of *-aʔ* and thus not predicted by previous analyses. In (6), this clause appears to simply describe the event of Coyote jabbing in the sagebrush with a stick in isolation; it is not clear that it is connected temporally or contrastively with any of the clauses surrounding it. In fact, the following clause in the text starts with *ʔudi* ‘and then’, indicating that the event in the following clause takes place after the event in this clause.

- (6) Context: Coyote is using sagebrush to build a fire under the rock where Lizard is hiding.

máʔaklu dá:bala t’ót’omuwaʔaʔ
maʔag-lu da:bal-a Ø-t’ot’om-uweʔ-aʔ
stick-INS sagebrush-LOC 3-jab-hence-DEP

‘He [Coyote] jabbed around in the sagebrush with a stick.’

(JW: 20)

Another example is in (7). Both clauses in this sentence contain an instance of *-aʔ*, but it is not clear that either of them uses *-aʔ* in a way that is predicted by Table 2. The first clause might plausibly be considered a temporal adjunct clause (e.g., ‘As the Coyote got kind of tired, he ran

away’), but it is not obvious that the second clause is connected temporally or contrastively with the following clause, which also starts with *ʔudi* ‘and then’. I counted both of these cases as matrix uses of *-aʔ*.¹⁰

- (7) Context: Coyote is chasing Lizard, but is not able to catch him.

ʔudi géwe gumyóʔlahé:šaʔ
 ʔ-ud-i gewe ʃ-gum-yoʔla-he:š-aʔ
 THEME-SEQ-IND coyote 3-REFL-tired-Q-DEP

Múʔšuwáʔaʔ
 ʃ-Mu-iʔiš-uweʔ-aʔ
 3-run-forward-hence-DEP

‘And then Coyote got kind of tired and ran away.’ (SJ: 15)

In sum, although there are some cases of dependent-marked clauses that are somewhat open to interpretation, there do appear to be uses of *-aʔ* in these texts that do not fit with the predictions of Table 2, and specifically where a dependent-marked clause seems to be used as a matrix clause.

4 Results

In the table below, I present for each text the number of matrix clauses with independent *-i* and dependent *-aʔ*, as a percentage of total main clauses. These numbers are instructive because matrix clauses are the only environment where a speaker in principle has a choice between using *-i* or *-aʔ*. As shown in Table 4, *-aʔ* is used more frequently than *-i* in main clauses in all the texts, i.e., across all speakers.¹¹

Table 4: Percentage of matrix clauses using *-i* and *-aʔ*

Text	Matrix clauses with <i>-i</i>	Matrix clauses with <i>-aʔ</i>
JW	9/29 = 31.0%	15/29 = 51.7%
BH	14/47 = 29.8%	30/47 = 63.8%
SA	2/22 = 9.1%	15/22 = 68.2%
SJ	1/12 = 8.3%	7/12 = 58.3%

¹⁰ The question morpheme *hé:š* seems to be used here as a hedge (‘kind of’), as the clause it appears in is not interpreted as a question.

¹¹ The percentages do not add up to 100% because some clauses contain a mood other than *-i* or *-aʔ*.

Since Bochnak and Hanink (2022) predict the percentage of matrix clauses with $-a?$ to be zero, non-zero values in the third column of Table 4 are problematic for their analysis. As it turns out, dependent-marked clauses are used in matrix clauses more than half the time by all speakers. This finding corroborates Jacobsen’s (1964) comments that $-a?$ in texts is used as a “narrative tense” of sorts.

5 Discussion

At this time, I see two possible ways to think about this data (there could also be others). First, we could still try to push the analysis of Bochnak and Hanink (2022) onto these apparently matrix uses of the dependent mood. Under such a view, the dependent-marked clauses that don’t seem to fit into the uses described in Table 2 would still be considered adjunct clauses, and are used in narratives as a general clause chaining strategy. An account along these lines is already suggested by Bochnak and Hanink (2022), and is compatible with their semantic analysis of $-a?$, whose only semantic content is conjunction as shown in (8). In (8), the dependent mood can conjoin properties of the same type (α stands in for elements of type e , i or s ; which one is chosen depends on the syntactic height of the adjunction site of the dependent-marked clause). But crucially, dependent $-a?$ does not directly lexicalize the temporal or contrastive readings, so this semantics is in principle compatible with the use of $-a?$ as a general clause chaining mechanism without any other semantic import.

$$(8) \quad \llbracket -a? \rrbracket = \lambda P_{\langle \alpha, t \rangle} \lambda Q_{\langle \alpha, t \rangle} \lambda x_{\alpha} [P(x) \ \& \ Q(x)]$$

A second possible avenue would be to take the suggestion by Jacobsen (1964) more seriously that there is something special about narratives, which is signalled by the widespread use of the dependent mood. On this kind of view, this special use of the dependent mood could be thought of as part of a wider phenomenon that TAM categories across languages can often have what appear to be “non-canonical” uses in narrative discourse. Some familiar examples include the narrative present in English and the reportative subjunctive in German.

In the so-called “narrative present” in English, the discourse is not anchored to the speech time but rather to the time that the narrative takes place (e.g., Anand and Toosarvandani 2018). An example of the English narrative present is given in (9).

- (9) Mr. Tulkinghorn takes out his papers, asks permission to place them on a golden talisman of a table at my Lady's elbow, puts on his spectacles, and begins to read by the light of a shaded lamp.

(Dickens, *Bleak House*, cited by Anand and Toosarvandani 2018)

It is not clear to me that the Wá·šiw dependent mood involves shifted temporal anchoring in the same way that the English narrative present shifts the anchoring of the present tense in narratives. For this reason, I will not consider this comparison any further here. Instead, I would like to suggest that the distribution of the Wá·šiw dependent mood bears some similarities with the German reportative subjunctive.

The reportative subjunctive (Konjunktiv I) in German often appears in the complement clause of a verb of saying. It can also appear in a matrix clause, and when it does, it leads to the interpretation that the content of that clause is being reported by the speaker as what someone else has said. For instance in (10), the Konjunktiv I is used in the first sentence in the complement clause of the verb *sagte* 'say'. The use of the Konjunktiv I in the following matrix clause leads to the inference that the content of this clause is also reporting what the subject of the first clause said. Although the second sentence is not syntactically subordinate to the verb of saying in the first sentence, there is a sense in which it is interpreted as if it were.

- (10) Er sagte, sie sei schön. Sie habe grüne
he say.PAST.INDIC she be.KONJI pretty she has.KONJI green

Augen.
eyes

'He said she is pretty. She has green eyes (he said).'

(adapted from Schlenker 2005)

Like the German Konjunktiv I, the dependent mood in Wá·šiw also appears in the complements of the verb *-i:d* 'say'. However, the distribution of the dependent mood is much wider than the Konjunktiv I; the former also appears in complements of other non-factive attitudes such as *-hamu* 'think', and also in temporal and concessive adjunct clauses. Nevertheless, there is a sense in which it may make sense to think of narratives such as the ones used in this study as being reportative, since *Coyote and Lizard* is a folklore tale passed down orally through generations. In fact, two of the four versions of the texts studied here (those from John Wiger and Sylvia Andrews) conclude with a verb of saying, where the story-

teller explains that they are recounting the story as it was told to them. The example from Sylvia Andrews is given in (11).¹²

(11) Context: Lizard is victorious over Coyote.

ʔudiš	t'ánu	píteliduŋ	lá:dugabaʔ
ʔ-ud-i-š	t'anu	píteliʔ-duŋ	le-a:duʔ-gab-aʔ
THEME-SEQ-IND-SR	person	lizard-like	1-hand-DIST.FUT-DEP
ʔítluliʔišgeduŋ		ʔéluliyaʔ	
ʔ- i d-luliʔ-i-š-ge-duŋ		ʔ-eʔ-luli-aʔ	
3- say -DIST.PAST-IND-SR-OBJ.REL-like		3-be-DIST.PAST-DEP	
ʔí:demelʔgi			
ʔ- i :d-emelʔ-i-gi			
3- say -TRAD-IND-SUBJ.REL			

‘Thus people were to have hands like Lizard’s, it was said and so it was long ago, the tradition says.’ (SA: 20)

In (11), we see the complements of the verb forms with *-i:d* ‘say’ contain the dependent *-aʔ*, as expected. Perhaps, then, we can understand the whole narrative as being implicitly interpreted as if it was subordinate to a verb of saying, and the near ubiquity of the dependent mood *-aʔ* as indicating just that. Note this (kernel of a) theory is not saying these uses of *-aʔ* are actually syntactically subordinate to a verb of saying in the narrative. Indeed, two out of the four texts considered here do not end with a verb of saying like (11). Rather, the idea is that the dependent mood in matrix clauses in Wá-šiw narratives is behaving like the Konjunktiv I in German — when it appears in a matrix clause, it is interpreted *as if* it is occurring subordinate to a verb of saying. I also do not wish to commit to saying that the dependent mood *conveys* reported speech, since it appears in many environments where reported speech is not implicated.

While I think the analysis sketched here has some plausibility, there are some challenges that would need to be worked out in more detail.

¹² This is actually the penultimate sentence of the text. The speaker ends the text with the sentence:

(i) Context: Immediately following (11).

díʔŋaŋ	ke	diYá:mle
díʔŋaŋ	ke	di-Yá:m-le
finish	TOP	1-tell-REDUND

‘That’s all I’ll say.’

(SA: 21)

The first is that, as observed in Table 4, there are still a minority of matrix clauses in the texts that contain the independent mood *-i*, which do not fall under the idea that the text as a whole is interpreted as reportative. A second issue is defining what counts as a narrative for the purposes of the matrix use of the dependent mood. Since not all versions of *Coyote and Lizard* end with a passage along the lines of (11), this kind of passage cannot be used as a signpost for delineating what counts as a narrative. Since these kinds of folklore narratives involve a monologue by a speaker, perhaps the monologue itself is enough to delineate the narrative and when we should expect to find these uses of the dependent mood. An examination of more texts may be useful in giving this question more traction.

For now, I hope to have made the case that the use of the dependent mood *-a?* in Wá·šiw narratives is a phenomenon that deserves further research to fully understand its use and meaning contribution. (Perhaps you have a clever idea, Hotze?)

References

- Anand, Pranav, and Maziar Toosarvandani. 2018. No explanation for the historical present: Temporal sequencing and discourse. In *Proceedings of Sinn und Bedeutung 22*, ed. Uli Sauerland and Stephanie Solt, 73–90. Berlin: Leibniz-Zentrum für allgemeine Sprachwissenschaft.
- Arregi, Karlos, and Emily Hanink. 2022. Switch reference as index agreement. *Natural Language and Linguistic Theory* 40:651–702.
- Bochnak, M. Ryan. 2016. Past time reference in a language with optional tense. *Linguistics and Philosophy* 39:247–294.
- Bochnak, M. Ryan. 2023. Some aspects of aspect in Wá·šiw. *International Journal of American Linguistics* 89:1–38.
- Bochnak, M. Ryan, and Emily Hanink. 2022. Clausal embedding in Washo: Complementation vs. modification. *Natural Language and Linguistic Theory* 40:979–1022.
- Evans, Nicholas. 2007. Insubordination and its uses. In *Finiteness: Theoretical and empirical foundations*, ed. Irina Nicolaeva, 366–431. Oxford: Oxford University Press.
- Jacobsen, William H. 1964. A grammar of the Washo language. Doctoral Dissertation, University of California, Berkeley. Available online at <http://escholarship.org/uc/item/52c6q7hg>.

- Portner, Paul. 2018. *Mood*. New York: Oxford University Press.
- Reisinger, Daniel, Lisa Matthewson, and Hotze Rullmann. 2022. Using corpora to investigate modal-temporal interactions. In *Modality in under-described languages: Methods and insights*, ed. Jozina Vander Klok, Núbia Ferreira Rech, and Simone Guessier, 141–190. Berlin: Mouton de Gruyter.
- Rullmann, Hotze, and Lisa Matthewson. 2018. Towards a theory of modal-temporal interaction. *Language* 94:281–331.
- Schlenker, Philippe. 2005. The lazy Frenchman’s approach to the subjunctive. In *Romance languages and linguistic theory 2003*, ed. Twan Geerts, Ivo van Ginneken, and Haike Jacobs, 269–309. Benjamins.

A second-last position clitic in Sm'algyax: a puzzle

COLIN BROWN

University of California, Los Angeles

HENRY DAVIS

University of British Columbia

1 Introduction

1.1 Preamble

We're very happy to be included in this volume celebrating Hotze and his work; one of us is a long-time colleague and some time co-author of his, and the other a past undergraduate student in his semantics class.

Our contribution comes in two parts. In the first, we present a particularly puzzling set of data involving the *wh*-clitic $=(d)u$ in Sm'algyax (Coast Tsimshian); in the second, we propose a solution.

The reader (and Hotze himself) might wonder how work on the morphosyntax of clitics in a Tsimshianic language relates to his own research, which — while it encompasses a wide variety of topics and languages — has never, as far as we are aware, touched on either of these areas. The answer is that though the subject matter might not relate directly to his own work, Hotze's influence can be felt in several ways in our contribution. First, though his background is in formal semantics, he has always paid close attention to the empirical details of natural language, which has made him for many years an ideal ally and resource for fieldworkers such as ourselves working on less-studied languages. Second, he has always appreciated puzzles and solutions (as evidenced by his strong support for the NACLO competition over the years). And third, an aspiration of ours which Hotze's work seems to embody effortlessly (though we know that a great deal of effort goes into it) is to take complex and apparently confusing data, reduce it to its essence, and offer a solution that in retrospect — but only in retrospect — seems intuitively correct.

We hope Hotze enjoys our attempt to emulate him!

1.2 Introduction to the puzzle

The Tsimshianic languages are known for their complex and varied clitic systems (see e.g. Mulder and Sellers 2010). In this squib, we examine a particularly challenging and intriguing clitic in Sm'algyax (Maritime Tsimshianic, ISO 639-3: *tsi*): the *wh*-clitic $=(d)u$. Our contributions are as follows: (i) we outline the syntactic distribution of the *wh*-clitic and suggest that it occupies a position high in the syntactic superstructure, taking an interrogative CP as its complement (Section 2); (ii) we show that it is phonologically an enclitic (Section 3); and (iii) we present a linearization puzzle associated with the *wh*-particle: sometimes it appears in a left-peripheral position encliticized to a *wh*-expression, and sometimes it appears in a clause-internal position encliticized to the predicate or a DP element (Section 4). The data and generalizations presented here lay the groundwork for Davis and Brown (this volume), which puts forth an analysis of the *wh*-clitic as a second-last position (penultimate) clitic.

2 The syntax of content questions and the *wh*-clitic

In this section, we describe content (*wh*-) question formation and outline the syntactic distribution of the *wh*-question marker $=(d)u$; we show that it is a root-level clitic that is restricted to content questions. It is neither a marker of clause type nor intrinsically associated with *wh*-expressions. Based on its behaviour and distribution, we suggest that the *wh*-clitic is a marker of interrogative illocutionary mood and is base generated above CP.

2.1 The structure of content questions

Content questions in Sm'algyax are characterized by the appearance of a *wh*-expression in clause-initial position together with extraction morphology that indicates the grammatical role of the extracted element, distinguishing between S (the subject of an intransitive predicate), A (the subject of a transitive verb), O (the object of a transitive verb), and others (adjuncts and oblique arguments). We illustrate this extraction morphology in the examples below. Content questions are additionally marked by the presence of the *wh*-clitic $=(d)u$, which appears in all the examples

below.^{1,2}

- (1) Naayu sis'aaxsit?
 naa=**du** sis'aaxs-**it** ___
 who=**Q** laugh-**sx**
 'Who laughed?' S-extraction
- (2) Goyu gaba gyet?
 goo=**du** gap-**i**=a gyet ___
 what=**Q** eat-**TR-3.II**=**CN** person
 'What do the people eat?' O-extraction
- (3) Naayu int gaba ts'ik'aaws?
 naa=**du** **in**=**t** gap=**a** ___ ts'ik'aaws
 who=**Q** **AX**=**3.I** eat=**CN** split.salmon
 'Who eats split dried salmon?' A-extraction

¹ Content questions, relative clauses, and focus fronting constructions are all marked with the same extraction morphology. Only content questions are marked with the *wh*-clitic. See Brown (2024) for a detailed description of the morphosyntax of extraction in Sm'algyax.

² Sm'algyax, also known as Coast Tsimshian or the Ts'msyen language, is spoken along the coast of Northern British Columbia, and on the island of Metlakatla, Alaska. All uncited examples come from elicitations with Velna Nelson, Ellen Mason (Txałgiiw/Hartley Bay), and Beatrice Robinson (Gitx-aala/Kitkatla). Linguistic examples are given in a four-line format: the top line is given in the Sm'algyax community orthography (Dunn 1978), the second line is presented in the same orthography, but indicates morpheme breaks — word-level morphophonological processes such as obstruent voicing before vowels are not marked at this level. The third line provides grammatical category labels, and the final line provides an English translation. Abbreviations for linguistic glosses are as follows: 1 = first person, 2 = second person, 3 = third person, AX = agent extraction morpheme, CN = common noun connective, COMP = complementizer, FOC = focus, I = series I clitic, II = series II suffix, III = series III pronoun, IRR = irrealis, NEG = negative, OBL = oblique, PASS = passive, PFV = perfective, PL = plural, PN = proper noun connective, POSS = possessive, PREP = preposition, PROSP = prospective, Q = question particle, REAS = reason subordinator, REL = relative, SG = singular, SX = subject extraction morpheme, T = "Big T" verbal morpheme, TR = transitive, VER = verum.

- (4) Ndeyu wil sa oksga l̥gwoomlk?
 ndeh=**du wil** sa=oks-k=a l̥gwoomlk ____
 where=**Q COMP** off=fall-PASS=CN child
 ‘Where did the child fall?’ Adjunct-extraction

Following earlier work on \bar{A} -movement in Tsimshianic (Davis and Brown 2011; Davis and Nederveen 2021), we adopt the hypothesis that there are two types of *wh*-questions, characterized by Davis and Brown as cases of “direct” versus “indirect” movement. Direct movement proceeds as in English: a *wh*-expression undergoes \bar{A} -movement to the left periphery. Indirect movement structures feature a predicative *wh*-expression that is base generated in initial position and takes a DP as its argument (typically a headless relative clause). Though the surface realization of direct and indirect movement is often identical, there is one construction in Sm'algyax that unambiguously signals the indirect movement structure: content questions featuring the relative pronoun *gu*. Originally, *gu* was likely a reduced form of the *wh*-expression *goo* ‘what’, but in the contemporary language it is not a question word. Instead, it introduces a relative clause, as shown below with a headed relative clause in (5) and a headless relative clause in (6).

- (5) Wilaayu hana'a gu sis'aaxsit.
 wilaay-i-u=a hana'a=a [**gu** sis'aaxs-it ____]
 know-TR-1SG.II=CN woman=CN REL laugh-SX
 ‘I know the woman that laughed.’

- (6) Gabu gu nah dzabn.
 gap-i-u=a [**gu** nah dzap-i-n ____]
 eat-TR-1SG.II REL PFV make-TR-2SG.II
 ‘I ate what you made.’

Gu may also appear in *wh*-questions, as shown in (7) below.

- (11) Akandi aap'ax ndeł
 aka=n=di aap'ax[-t] [ndeh(*=du)=ł
 NEG=1SG.I=FOC remember[-3.II] where(*=Q)=IRR.CN
 habit.
 hap-i-t ___]
 PL:go-TR-3.II
 'I don't remember where they went.'

Wh-expressions also appear in a number of non-interrogative contexts, including as indefinite/indeterminate nouns (12)–(14), in headless relative clauses (15), and in exclamatives (16). The *wh*-clitic is strictly prohibited from appearing in any of these constructions.

- (12) Ła'a ligi goo haasgu.
 ła'a=a ligi **goo**(*=du)=a haas-k-u
 bite=CN LIGI **what**(*=Q)=CN dog-PASS-1SG.II
 'Something bit my dog.'

- (13) Nah niidzu ligit naa.
 nah niits-u ligi=t **naa**(*=du)
 PFV see-1SG.II LIGI=PN **who**(*=Q)
 'I saw someone.'

- (14) Dm małdu txa'nii goo da k'wan.
 dm mał-t-i-u txa'nii **goo**(*=du) da k'wan
 PROSP tell-T-TR-1SG.II all **what**(*=Q) PREP 2SG.OBL
 'I will tell you everything.'

- (15) Waayu naa dmt in dzaba
 Waa-i-u [**naa**(*=du) dm=t in dzap[-t]=a ___
 find-TR-1SG.II **who**(*=Q) PROSP=3.I AX do[-3.II]=CN
 ts'ikts'igu.
 ts'ikts'ik-u]
 car-1SG.II
 'I found someone who will fix my car.' Lit. 'I found **who** will fix my car.'

- (19) Context: You're talking to a friend who returned from a baking exchange:

Goyu nah gabn ada naat nah int
 [goo=**du** nah gap-i-n ___] [ada naa=t nah in=t
 what=**Q** PFV eat-TR-2SG.II and who=IRR.CN PFV AX=3.I
 dzapt?
 dzap-t ___]
 make-3.II
 'What did you eat and who made it?'

- (20) [_{MoodP} [_{ConjP} [_{CP} WH ... [_{IP} ...]]] [_{Conj'} & [_{CP} WH ... [_{IP} ...]]]] =(*d*)*u*]

The *wh*-clitic cannot scope over the second conjunct from its surface position inside the first conjunct. Furthermore, syntactic movement into a c-commanding position would violate the Coordinate Structure Constraint (Ross 1967). The only other possibility — which we will adopt — is that =(*d*)*u* is base-generated in a c-commanding position above both conjuncts (as in (20)) and is positioned inside the first conjunct post-syntactically.

To conclude this section, the *wh*-clitic =(*d*)*u* only appears in root *wh*-questions; it is absent from embedded questions and incompatible with non-interrogative uses of *wh*-expressions. We suggest, based on this distribution, that the *wh*-clitic is generated in an illocutionary MoodP above CP and selects an interrogative CP complement.

3 Phonology

The goal of this section is to show that, phonologically, =(*d*)*u* must lean on some phrase to its left. We present four pieces of evidence for this claim. First of all, the *wh*-clitic never appears in initial position: this follows straightforwardly from its enclitic status.

- (21) *U/Yu/Dunaa liimit?
du=naa=a liimi-it?
Q=who=CN sing-SX
 Intended: 'Who sang?'

The second piece of evidence comes from the interaction of $=(d)u$ with the determiner-like elements known as “connectives” in the literature on Tsimshianic (see Davis 2018 and references therein). Connectives are syntactically associated with a nominal element to their right, but phonologically encliticize to a phrase to their left (Mulder and Sellers 2010). For example, in (22) below, the proper noun (PN) connective $=t$ introduces the pronoun *'nüün* that appears to its right, but encliticizes to the sequence of the *wh*-expression *naa* plus the *wh*-clitic $=(d)u$ that appears to its left.⁴ If $=t$ is encliticized to the *wh*-phrase, and $=(d)u$ precedes it, $=(d)u$ must also be an enclitic.

- (22) Naayu 'nüün? (not: **naa=t=(d)u 'nüün*)
 naa=**du** [=t 'nüün]
 who=Q =PN 2SG.III
 ‘Who are you?’

Third, we observe contextual allomorphy effects that are triggered when $=(d)u$ encliticizes to a *wh*-word. When the *wh*-clitic immediately follows a *wh*-word, it optionally surfaces as either [ju] or [du]:

- | | |
|---|---|
| <p>(23) Naayu baat?
 naa=du baa-it
 who=Q run-SX
 ‘Who ran?’</p> | <p>(24) Naadu baat?
 naa=du baa-it
 who=Q run-SX
 ‘Who ran?’</p> |
|---|---|

However, when the *wh*-clitic follows a non-*wh* word, it obligatorily surfaces as [du]:⁵

- (25) Goł wils liimidu dm yaatm?
 goo=l wils liimi=**du** dm yaat-m
 what=IRR.CN kind song=Q PROSP tell/sing-1PL.II
 (not: **gol wils liimiyu*)
 ‘What kind of song will we sing?’ (SLLTD)

⁴ Proper noun connectives (also known as “determinate” connectives in the Tsimshianic literature) introduce independent (Series III) pronouns as well as proper nouns in Sm'algyax.

⁵ We address the linear position of $=du$ in such cases in Section 4 immediately below.

Assuming that contextual allomorphy of this type requires not only adjacency but phonological integration with the preceding word, these data provide another argument that $=(d)u$ must be enclitic to the *wh*-word.⁶

A fourth piece of evidence comes from deletion of final /t/ when immediately followed by $=du$. Recall that when the *wh*-clitic follows a non-*wh* element, it always surfaces as [du]. In (26) below, we see a *wh*-clitic following the possessive *wh*-phrase *naal naboodit* ‘whose boat’, which ends with the *-it* suffix characteristic of possessor extraction. The surface form shows that sequences of *-it* and $=du$ reduce to [idu], rather than surfacing as [itdu], suggesting that $=du$ is phonologically integrated enough with the phrase to its left to condition deletion.⁷

- (26) Naal naboodidu giikt?
 naa=ɪ na=boot-**it=du** giik-t
 who=IRR.CN POSS=boat-SX=Q buy-3.II
 (not: **nabooditdu*)
 ‘Whose boat did she buy?’

The ban on the *wh*-clitic appearing in initial position, as well as the contextual allomorphy and deletion facts associated with the element to the left of the *wh*-clitic all point to the same conclusion: $=du$ is phonologically an enclitic.

4 Linearization

So far, nearly all the cases of *wh*-questions we have seen show $=du$ in the clausal second position, immediately following a *wh*-expression.⁸ This is compatible with both its syntactic position, as expounded in Section 2, and its phonological properties, as described in Section 3: it occupies a

⁶ Since the appearance of the allomorph [ju] is a predictable (albeit optional) consequence of the *wh*-clitic encliticizing to a *wh*-word, from now on, we write the morpheme $=(d)u$ simply as $=du$.

⁷ Another possible explanation for this alternation is that the onset of the *wh*-clitic undergoes deletion, and the coda of the *-it* morpheme undergoes voicing before [u] resulting in the attested surface form of [idu]. Either option supports the central claim that $=du$ is phonologically integrated with the phrase to its left.

⁸ The exception is (25) above.

position high in a root clause (with CP as its sister), and must attach to a phrase to its left, as a prosodically dependent enclitic.

However, in spite of what might appear to be the case from the examples provided so far, =*du* is not at all confined to second position. In fact, it turns out to be a typologically unusual second *last* position (penultimate) clitic, as we now show by giving a more complete picture of its distribution.

We can characterize this distribution as falling into three patterns. The first is where =*du* occurs attached to a *wh*-phrase at the left periphery of the clause, as schematized in (27).

(27) [WH=**du** [...]] *Wh-placement*

This is the main environment where we have encountered =*du* so far. See examples (5), (7), etc.

Second, =*du* encliticizes to the inflected predicate (typically but not exclusively a verb) following a *wh*-phrase and preceding an argument DP in any of S, A, or O function, as schematized in (28).

(28) [WH...[V=**du** DP_{S/A/O}]] *Predicate placement*

Examples of this pattern are given below. In (29), O is extracted and =*du* precedes A;⁹ in (30), A is extracted and =*du* precedes O, and in (31) and (32), an adjunct is extracted and =*du* precedes S.

(29) Goł gabidu gyet?
 goo=t gap-i-t=**du**=a gyet
 what=IRR.CN eat-TR-3.II=Q=CN person
 ‘What do the people eat?’

(30) Naat int gapdu ts'ik'aaws?
 naa=t in=t gap-t=**du**=a ts'ik'aaws
 who=IRR.CN AX=3.I eat-3.II=Q=CN split.salmon
 ‘Who ate the split salmon?’

⁹ The example in (29), which exemplifies *predicate placement*, forms a minimal pair with (2), which exemplifies *wh-placement*.

- (31) Dzindał dm 'ap yaltgidut Norman?
 dzindaa=ł dm 'ap yaltk-t=**du**=t Norman
 IRR.when=IRR.CN PROSP VER return-3.II=Q=PN Norman
 'When is Norman really coming back?' (Sasama 2001:64)

- (32) Ndał mi wil gyiikdu ngwüda'atsn?
 ndaa=ł mi wil gyiik-t=**du**=a n-gwüda'ats-n
 where=IRR.CN 2SG.II COMP buy-3.II=Q=CN POSS-coat-2SG.II
 'Where did you buy your coat?'

Third, =*du* attaches to the end of a DP in A function in a WH-V-A-O configuration, as schematized in (33) below.

- (33) [WH V DP_A=**du** DP_O] *Argument placement*

Such examples involve the *wh*-extraction of either an adjunct, as in (34), or an oblique argument, as in (35).

- (34) Goł gant dzapdit Meelidu ts'ikts'ik?
 goo=ł gan=t dzap-t=t Meeli=**du**=a ts'ikts'ik
 what=IRR.CN REAS=3.I do-3.II=PN Mary=Q=CN car
 'Why did Mary fix the car?'

- (35) Goł naht giindit Michaeldut Henry?
 goo=ł nah=t giin-t=t Michael=**du**=t Henry
 what=IRR.CN PFV=3.I give-3.II=PN Michael=Q=PN Henry
 'What did Michael give Henry?'

These three patterns (*wh*-placement, *predicate placement*, and *argument placement*) constitute the core distribution of =*du*. Note that *predicate placement* and *argument placement* are in free variation with *wh*-placement, but they are never in free variation with each other.

Before we put forth an explanation for this distribution, it is important to note elements which do not affect the position of =*du*. To start with, functional heads such as complementizers, tense/aspect markers, and subject clitics have no effect on its placement.

- (36) Ndeł nam wil niisdu ol?
 ndeh=ł **nah=m wil** niis-t=**du** ol
 where=IRR.CN **PFV=2.I COMP** see-3.II=Q bear
 ‘Where did you see the bear?’

- (37) Ndeł wil sa oksgadu lgwoomlk?
 ndeh=ł **wil** sa=oks-k-t=**du**=a lgwoomlk
 where=IRR.CN **COMP** off=fall-PASS-3.II=Q=CN child
 ‘Where did the child fall?’

These examples show that the linear position of =*du* is at least partially insensitive to syntactic structure: even though, as shown in Section 2, =*du* is base-generated at the very top of a root clause, in these cases it surfaces inside its CP complement. On a syntactic account, lowering would be required to derive its surface position; we take it that this is not a viable option.

Second, the presence of PPs (including oblique arguments as well as adjuncts) does not affect placement of the *wh*-clitic. This means that examples such as (38) and (39) below involve predicate placement: the bracketed PPs introduced by the preposition *da* have no effect on the linear position of =*du*, which ends up encliticized to the verb, followed by an argument DP (rather than encliticizing to the argument DP, as would be expected if the PP counted for clitic placement).

- (38) Goł ky'ilamdu 'yuuta da haas?
 goo=ł ky'ilam-i-t=**du** 'yuuta [da=a haas]
 what=IRR.CN give-TR-3.II=Q man PREP=CN dog
 ‘What did the man give the dog?’

- (39) Naal nah habooltidut Dzon asda
 naa=ł nah habool-t-i-t=**du**=t Dzon [asda
 who=IRR.CN PFV look.after-T-TR-3.II=Q=PN John PREP
 gits'iipda?
 gits'iipda]
 yesterday
 ‘Who did John look after yesterday?’

The same is true of CPs, as illustrated by the long-range *wh*-dependencies in (40)–(41):

- (40) Ndeł małdidut Betty gooys Meeli?
 ndeh=ł mał-t-i-t=**du**=t Betty [go-i[-t]=s Meeli]
 where=IRR.CN say-T-TR=Q=PN Betty go-TR-3.II=PN Mary
 ‘Where did Betty say Mary went?’

- (41) Goł ha'ligoodut Bettyt giindit
 goo=ł ha'ligoot-t=**du**=t Betty [=t giin-t=t
 what=IRR.CN think-3.II=Q=PN Betty =3.I give-3.II=PN
 Michaelt Henry?
 Michael=t Henry]
 Michael=PN Henry
 ‘What does Betty think Michael gave Henry?’

In these examples, =*du* again encliticizes to the verb rather than to the subject, as would be expected were the embedded CP to be treated like a DP object. In other words, these are again cases of predicate placement rather than argument placement: the embedded CP has no effect on clitic linearization. It is also important to note that the argument-adjunct distinction is not at play here: the embedded CPs in (40)–(41) are complements of the matrix predicate (and as such allow *wh*-extraction), yet as *non-DP* complements, they are invisible for the purposes of clitic linearization.

The relevant generalization covering all these cases is as follows:

- (42) *Only the predicate and its DP arguments count for the linearization of =du*

This shows that linearization of =*du* is partially *sensitive* to syntactic structure, in that it pays selective attention to the categorial signature of potential hosts.

5 Interim conclusion

This squib has introduced the *wh*-clitic =*du*, and outlined (i) its syntactic position in MoodP above CP; (ii) its phonological status as an enclitic; and (iii) the three linear positions it occupies (following a *wh*-phrase, following the predicate, and following an object DP). In Davis and Brown (this volume), we sketch a unified account of these three linear positions.

References

- Beck, Sigrid. 2006. Intervention effects follow from focus interpretations. *Natural Language Semantics* 14:1–56.
- Brown, Colin. 2024. Questions and their relatives in Sm'algyax. To appear in the *International Journal of American Linguistics*, 90(3).
- Cable, Seth. 2007. The grammar of Q: Q-particles and the nature of *wh*-fronting, as revealed by the *wh*-questions of Tlingit. Doctoral Dissertation, MIT, Cambridge, MA.
- Cable, Seth. 2010. Against the existence of pied-piping: evidence from Tlingit. *Linguistic Inquiry* 41:563–594.
- Cinque, Guglielmo. 1999. *Adverbs and functional heads: A cross-linguistic perspective*. Oxford University Press, Oxford.
- Davis, Henry. 2018. Only connect! determiners, case and agreement in Tsimshianic. *International Journal of American Linguistics* 84:471–511.
- Davis, Henry, and Colin Brown. This volume. A second-last position clitic in Sm'algyax: a solution.
- Davis, Henry, and Jason Brown. 2011. On A'-dependencies in Gitksan. In *Papers for the 46th International Conference on Salish and Neighbouring Languages*, 43–80. UBC Working Papers in Linguistics.
- Davis, Henry, and Sander Nederveen. 2021. Hli, focus and relativization in Gitksan. In *Papers for the 56th International Conference on Salish and Neighbouring Languages*, ed. D. K. E. Reisinger, Hannah Green, Laura Griffin, Marianne Huijsmans, Gloria Mellesmoen, and Bailey Trotter., 1–32. UBC Working Papers in Linguistics.

- Dunn, John Asher. 1978. A practical dictionary of the Coast Tsimshian language. Canadian Ethnology Service paper 42, Ottawa, ON.
- Kotek, Hadas. 2014. Composing questions. Doctoral Dissertation, MIT, Cambridge, MA.
- Kratzer, Angelika, and Junko Shimoyama. 2002. Indeterminate pronouns: The view from Japanese. In *The 3rd Tokyo Conference on Psycholinguistics*, ed. Yukio Otsu. Tokyo: Hituzi Syobo.
- Mulder, Jean, and Holly Sellers. 2010. Classifying clitics in Sm'algyax: Approaching theory from the field. In *Fieldwork and linguistic documentation in indigenous languages of the Americas*, ed. Andrea L. Berez, Jean Mulder, and Daisy Rosenblum, 2, 33–56. URL <http://nflrc.hawaii.edu/ldc/>, Language Documentation and Conservation Special Publication.
- Portner, Paul. 2018. *Mood*. Oxford University Press.
- Ross, John Robert. 1967. Constraints on variables in syntax. Doctoral Dissertation, MIT.
- Sasama, Fumiko. 2001. A descriptive study of the Coast Tsimshian morphology. Doctoral Dissertation, Kyoto University.
- SLLTD. 2017. Sm'algyax Living Legacy Talking Dictionary. <https://www.webonary.org/smalgyax/> [Accessed: August 2023].
- Uegaki, Wataru. 2018. A unified semantics for the Japanese Q-particle ka in indefinites, questions and disjunctions. *Glossa: a journal of general linguistics* 3(1): 14:1–45.

Evidence for two types of future semantics by negation*

Sihwei Chen
Academia Sinica

1 Introduction

In English there is only one sentential negator *not*. In (1a), it is not immediately clear whether the negation scopes over or under the future auxiliary *will*. Compare (1a) with *It is not going to rain*. If we assume with Copley (2009) that *be going to* consists of a progressive over a prospective operator *woll*, which is realized as *go*, the order in which the negator precedes *go* suggests that it has scope over *go*. Does the scope relation of *will not* also follow the surface scope as in the case of *be not going to*?

In the literature, it is well-established that future statements do not exhibit scope in relation to negation (Copley 2009; Cariani & Santorio 2018). This property is demonstrated by the law of excluded middle: $p \vee \neg p$, meaning that p is either true or its negation is true. Future statements, much like non-future ones, adhere to this law, as seen in a comparison with modal statements involving negation. For example, (1a) and (1b) share identical truth conditions, while (2a) and (2b) do not. This implies that *will*, if considered a future modal, must be scopeless, distinguishing it from other modals.

* I had the privilege of taking three of Hotze's seminar courses, two consecutive ones from September 2011 to April 2012 (Modality and Tense) and another in the fall of 2013 (Number and Quantification), and serving as a teaching assistant for his undergraduate class on pragmatics. These experiences significantly enriched my understanding of semantics and pragmatics.

Hotze has an extraordinary talent for explaining complex concepts. His expertise in analyzing language facts through diverse linguistic interfaces also left a lasting impression on me. Having him on my dissertation committee was a further privilege. I vividly remember the moment he promptly agreed when I approached his office door, and his invaluable insights guided me through every stage of my research challenges. We later worked closely together on several projects, and I learned a lot from his knowledge and research methods. Hotze consistently fostered a positive atmosphere in our meetings, sharing insightful observations, comments, and humor. I would like to take this opportunity to express my gratitude. *sqasay ta' inbyaqan su!* (lit. 'Let us be joyful for your birth!').

- (1) a. It **will not** rain.
b. It is **not** the case that it **will** rain.
- (2) a. Sam **must not** lock the door.
b. It is **not** the case that Sam **must** lock the door.

There are two key assumptions in this comparison: (i) *will* lexically encodes quantification over possible worlds, and (ii) *will* takes scope over *not* in (1a), similar to how *must* takes scope over *not* in (2a). The question that arises is whether both assumptions hold. Cariani and Santorio (2018) challenge the first assumption. They argue that *will* is a modal that takes as argument a modal base pronoun but does not quantify over worlds; instead, it selects a unique world within the modal base, which by default is the evaluation world (due to certain closeness conditions). This analysis renders *will* semantically vacuous with respect to the world parameter and leads them to the desired result that *will not p* (1a) and *not will p* (1b) are equivalent.

However, Cariani and Santorio do not address the question of how the temporal semantics of *will* interacts with negation since it is not their focus. They assume that *will* existentially quantifies over future times, while noting that an alternative approach is an interval extending forward from the evaluation time. Even if we adopt their analysis that *will* is a vacuous modal (see their review of earlier approaches),¹ an existential operator can still potentially take scope above or below a sentential negator, the result of which would not be equally predicted by the postulation of an unbounded interval.

In this paper, I present data from two languages, Atayal and Mandarin, to demonstrate the necessity of distinguishing between the two types of forward-shifting semantics mentioned above. Atayal is an Austronesian language spoken in northern Taiwan, and the data presented here primarily come from my own fieldwork, unless stated otherwise. The Mandarin data exclusively pertain to the dialect spoken in Taiwan and are based on my own intuitions and consultations with native speakers. Atayal follows a predicate-initial word order, while Mandarin employs an SVO word order, akin to English. Both languages fall under the

¹ Alternatively, we can decompose *will* into a covert modal and a prospective aspect *woll*. In this analysis, the negation can take scope above *woll* but below the modal, thus avoiding the negation of universal quantification over worlds.

category of morphologically tenseless languages with overt future morphemes, which usually cannot be omitted in future contexts.

While it is not always easy to determine the scope between *will* and *not* in English, the interaction between future and negation becomes clear in languages where the negator is syntactically constrained to appear either above or below a future marker. In Section 2, I show that there are distinct syntactic negators in Atayal, with one taking scope over future modals and the other appearing below various circumstantial modals but not below the future modals. In Section 3, I propose an explanation for these patterns by assigning existential quantification over times to the future modals and an open interval to the circumstantial modals. This proposal is supported by Mandarin, which has an internal negator solely for negating the existence of events and cannot take scope under the future modal. Section 5 concludes with a semantic typology.

2 The interaction between standard negators and future-oriented modals in Atayal

2.1 External and internal negators

In Atayal, there are two standard negators that function similarly to sentential negation: *iyat* and *ini*'. Both negators are auxiliaries that attract bound pronouns (in the absence of another higher auxiliary), but they exhibit significant morphosyntactic differences. One of the most striking is the voice inflection of the verb following the negator: *iyat* requires that the following verb be in the indicative (which is unmarked in glosses) (3–4), while *ini*' requires it to be in the dependent (5) (i.e., *kita*' instead of *mita*').²

(3) a. **iyat p-qwalax** rihay 'nyal.
 NEG AV.FUT-rain week come.NMLZ
 'It will not rain next week.'

b. **iyat=nya'** **niq-un** qu hi' bzyuwak hiya'.
 NEG=3SG.ERG eat-PV.FUT ABS body boar EMP
 'He will not eat the pork.'

² Abbreviations that are not in the Leipzig Glossing Rules: ABIL, ability; AV, actor voice; CIRC, circumstantial; COS, change of state; CTF, counterfactual; CV, circumstantial voice; DEON, deontic; DEP, dependent; EMP, emphatic; EPIST, epistemic; E.PST, existential past; EXP, experiential; LV, locative voice; NAV, non-actor voice; NEC, necessity; POS, possibility; PRT, particle; PV, patient voice.

- (4) **iyat**=maku' **k<in>i'-an** ngasal qani.
 NEG=1SG.ERG live<E.PST>-LV house this
 'I have not/never lived in this house.'
- (5) **cyux** **ini'** **kita'** biru' qu hiya'
 PROG.DIST NEG see.AV.DEP book ABS 3SG.N
 'He is not reading books.'

A brief introduction to the voices of Atayal is needed here. Atayal has a typical Philippine-type voice system, i.e., each verb must be marked with one of the four voices that vary in the macro-thematic role of the subject (i.e., Actor Voice, Patient Voice, Locative Voice, and Circumstantial Voice). At the same time, the voice also varies with three mood groups that roughly correspond to what is called sentence mood: indicative, dependent, and hortative. Only in the indicative, but not in the dependent mood, can future and past affixes be present (i.e., *p-* in (3a) and *-in-* in (4)). The examples in (3) also illustrate a morphological peculiarity that in sentences that are not in AV, the future prefix *p-* disappears, making the sentence appear in the same form as non-future sentences (Chen 2018:279ff.); hereafter I refer to them as *p*_{AV} and \emptyset _{NAV}.

While the voice inflection of the main verbs after the negators may simply reflect their morphological difference, other distinctions suggest that the two negators have different syntactic positions. For example, *ini'* directly precedes the verb, so a freely distributed adverb cannot intervene between them, and *ini'* must follow an overt aspect marker (5), while *iyat* lacks these features. The examples in (3) to (5) establish the syntactic hierarchy of the two negators within the clause: *ini'* is below AspP, and *iyat* is above TP, where TP is conventionally above AspP.³ The resulting hierarchy is presented as (6).

- (6) *iyat* > TP > AspP > *ini'* > VP

Based on the above and other evidence, I assume that *iyat* negates the entire proposition by taking a TP of type $\langle s, t \rangle$ as an argument, while *ini'* negates the event denoted by the predicate by taking an eventuality of

³ The overt temporal morphemes that are supposed to occupy the tense head, *p*_{AV} and *-in-*, do not co-occur with aspect markers; sentences with aspect are usually morphologically tenseless.

type $\langle l, st \rangle$ (7); such a distinction is often called external and internal negation.⁴

- (7) a. $[[iyat]]^{g,c} = \lambda P_{\langle s,t \rangle} \lambda w. \neg[P(w)]$
 b. $[[ini']]^{g,c} = \lambda P_{\langle l,st \rangle} \lambda e \lambda w. \neg[P(e)(w)]$

2.2 Ways of interaction with future-oriented modals

As expected from their syntactic position, the two negators interact differently with epistemic and circumstantial modals, which are also in a higher and lower position, respectively (Chen 2018:425ff.). The examples in (8) show that epistemic modals asymmetrically precede both negators. In contrast, circumstantial modals may precede the internal negator *ini'* (9), but do not co-occur with the external negator *iyat* in either order (e.g., **iyat nway* ... or **nway iyat* ...). The same pattern holds for the counterfactual/irrealis marker *aki* 'would' (10), the habitual/generic marker *mutux*, and markers used specifically in purpose clauses or contexts of apprehensive/timitive modality (e.g., *teta'*/*tayta'* and *hala*).

- (8) a. **ki'a** **iyat** p-swal wah.
 EPIST.POS NEG AV.FUT-promise PRT
 'He might not agree.'
- b. **ki'a** **ini'** swayal qu Tali'.
 EPIST.POS NEG promise.AV.DEP NOM Tali'
 'Tali' might not have agreed.'
- (9) **nway=ta'** **ini'** p-qsya'-i kira' la.
 DEON.POS=1PL.ERG NEG CAUS-water-PV.NEG later.today COS
 'We don't need to water the vegetables today.' or 'We may not
 water the vegetables today.'

⁴ The terms 'external' and 'internal' simply refer to the sentence-external vs. -internal modification, and not to the semantic sense 'it is not the case that...' as in (1b) and the lack thereof. They are sometimes used depending on whether the negation negates a presupposition (i.e., a metalinguistic negation; cf. Horn 2001). However, the Atayal negation sentences in my data do not involve the cancelation of a presupposition, so it is less likely that *iyat* and *ini'* are specialized for metalinguistic negation.

- (10) *pung* *ke'=maku'* *ki.* ***aki***=*su'* ***ini'***
 listen.AV.IMP word=1SG.GEN PRT CTF=2SG.ABS NEG
ktakuy.
 fall.down.AV.DEP
 'You should listen to my words. You would not fall down.'

Interestingly, future modals (i.e., those used for prediction) do not behave either like epistemic or circumstantial modals in their interaction with the two negators. We have seen the future affix $p\text{-AV}/\emptyset_{\text{NAV}}$ necessarily after *iyat* in (3), and because of the requirement for a different mood, $p\text{-AV}/\emptyset_{\text{NAV}}$ does not co-occur with *ini'*. An auxiliary grammaticalized from the verb of going, which I refer to as *musa'*_{FUT}, can often alternate with $p\text{-AV}/\emptyset_{\text{NAV}}$ in affirmative contexts of prediction (Chen 2018:311ff.) (11), but unlike $p\text{-AV}/\emptyset_{\text{NAV}}$, *musa'* cannot co-occur with *iyat* in either order (12).

- (11) a. *ki'a* ***p***-*qwalax* *hazi'*.
 EPIST.POS AV.FUT-rain EPIST.POS
 'It might rain.' or 'It will possibly rain.'
- b. *kt-an* *kayal ga,* *hazi'* ***musa'*** *m-qwalax.*
 see-LV sky TOP EPIST.POS FUT AV-rain
 'It looks like it might rain.'

- (12) * {***iyat musa'*** / ***musa' iyat***} *m-qwalax rihay 'nyal.*
 NEG FUT FUT NEG AV-rain week come.NMLZ
 Intended: 'It will not rain this week.'

The incompatibility with external negation makes *musa'*_{FUT} resemble circumstantial modals. However, unlike circumstantial modals, e.g., (9), *musa'*_{FUT} is also incompatible with internal negation (13a). To render the reading 'will not', the future marker $p\text{-AV}/\emptyset_{\text{NAV}}$ must be used under external negation (13b).⁵

- (13) a. # ***musa' ini'*** *pawng-i* *k~kayal=su'*.
 FUT NEG listen-LV.DEP CV.NMLZ~say=2SG.GEN
 'Your words will not be heard.'

⁵ The sequence *musa' iyat* in (13b) does not contradict the generalization that *musa'*_{FUT} cannot co-occur with *iyat* (as in (12)). This is because in this case, 'must' functions as an epistemic modal and occupies a higher position, allowing it to be used with a present preajacent. As an epistemic modal, *musa'*_{EPIST} behaves similarly to the epistemic modal *ki'a* in (8a) by preceding *iyat*.

- b. *musa' iyat pawng-an kay'=su'.*
 EPIST NEG listen-LV.FUT word=2SG.GEN
 ‘Your words will probably not be heard.’
 Consultant: “They might mute your microphone.”

As the marking of the above data reveals, I suggest that (12) is ungrammatical and (13a) is infelicitous. The former is based on the observation that *iyat* cannot precede or scope over any modal in Atayal.⁶ However, the non-acceptance of (13a) is surprising when compared to other circumstantial modals, as in (9) and (10). Across languages, both types of modals are future-oriented (Condoravdi 2002), and future modals are often regarded as circumstantial. While some suggest ‘realis’ negation for *ini'* (e.g., Su 2004:66), its compatibility with circumstantial modals calls for a different explanation. In Section 3, I explore an explanation rooted in semantic (in)felicity arising from the distinct future semantics of *musa'*_{FUT} and the circumstantial modals.

Table 1 summarizes the discussion by listing the possible order and marking unattested co-occurrences for grammatical reasons as ‘N/A’. Focusing on circumstantial and future modals, we can identify three patterns A–C, highlighted in different colors. The main question here is why *musa'*_{FUT}, unlike the circumstantial modals, cannot take scope over *ini'*.

Table 1: Interaction between standard negators and modals in Atayal

	<i>iyat</i> ‘external NEG’	<i>ini'</i> ‘internal NEG’	Pattern
Epistemic	<i>ki'a</i> ‘might’	<i>ki'a</i> > <i>iyat</i>	<i>ki'a</i> > <i>ini'</i>
Circumst.	<i>nway</i> ‘can’	N/A	<i>nway</i> > <i>ini'</i>
	<i>aki</i> ‘would’	N/A	<i>aki</i> > <i>ini'</i>
Future	<i>musa'</i> _{FUT}	N/A	# <i>musa'</i> _{FUT} > <i>ini'</i>
	<i>p-_{AV}/Ø_{NAV}</i>	<i>iyat</i> > <i>p-_{AV}/Ø_{NAV}</i>	N/A

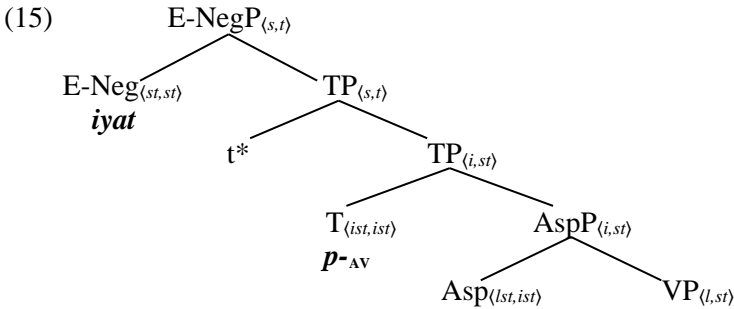
⁶ One possible explanation is that Atayal modals are grammaticalized into auxiliaries through the restructuring of complex clauses, such as conditionals or embeddings (Chen 2018:425ff.; Wu 2013:112ff.), and may occupy a similar syntactic position to external negation.

3 Proposal

As mentioned earlier, two types of forward-shifting semantics have been used in the literature: existential quantification and a right open interval. The latter is proposed in Abusch (1998) for English future modals and extended by Condoravdi (2002) to all circumstantial modals, which she argues are all future-oriented. I propose that in Atayal, the future modals $p\text{-AV}/\emptyset_{\text{NAV}}$ and $musa'_{\text{FUT}}$ encode the first option, while the circumstantial modals encode the second option. This proposal is illustrated by the formulas in (14), where following Condoravdi, I use $[t, _)$ to represent an interval that has t as its initial subinterval and extends to the end of time. I simplify the modal semantics in the circumstantial modal as MB_{Circ} , representing Kratzer's conversational background.

- (14) a. $\llbracket p\text{-AV}/\emptyset_{\text{NAV}}/musa'_{\text{FUT}} \rrbracket^{g,c} = \lambda P_{(i,st)} \lambda t \lambda w. \exists t' [t < t' \ \& \ P(t')(w)]$
 b. $\llbracket nway \rrbracket^{g,c} = \lambda P_{(i,st)} \lambda t \lambda w. \exists w' [w' \in MB_{\text{Circ}}(w, t)] \ \& \ P(w')([t, _))]$

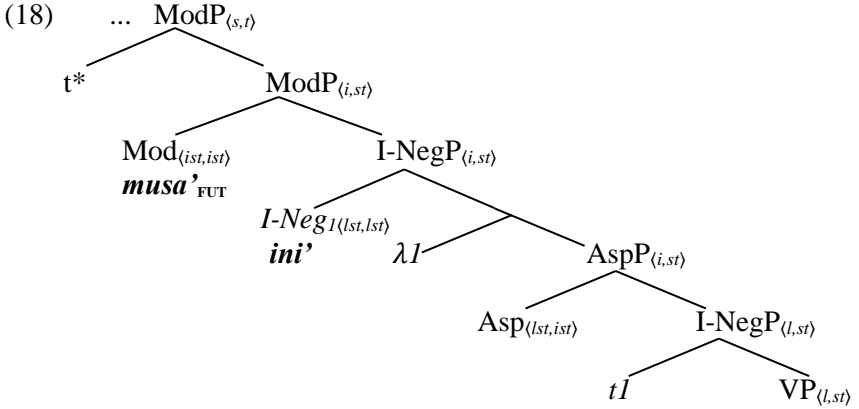
Let us start with pattern C in Table 1. In (15), we have the structure of (3a) with the lexical entry of VP and AspP in (16a, b). Here the external negator *iyat* scopes over the existential quantifier in the future prefix $p\text{-AV}$, resulting in the desired meaning: ‘there is no time in the next week when it rains’ (17).



- (16) a. $\llbracket VP \rrbracket^{g,c} = \lambda e \lambda w. \lambda P(e)(w)$
 b. $\llbracket PFV \rrbracket^{g,c} = \lambda P \lambda t \lambda w. \exists e [P(t)(w) \ \& \ \tau(e) \subseteq t]$

- (17) $\llbracket (4a) \rrbracket^{g,c} = \llbracket iyat \rrbracket^{g,c} (\llbracket \lambda w. \exists t' [t^* < t' \ \& \ t' \subseteq \text{next week} \ \& \ \exists e [\text{rain}(e)(w) \ \& \ \tau(e) \subseteq t'] \rrbracket^{g,c}) = \lambda w. \neg \exists t' [t^* < t' \ \& \ t' \subseteq \text{next week} \ \& \ \exists e [\text{rain}(e)(w) \ \& \ \tau(e) \subseteq t'] \rrbracket$

Pattern B is exemplified by the infelicitous example in (12), and its structure is provided in (18).



Despite the internal negator *ini'* being adjacent to VP (5) and its denotation in (7b), I analyze that it undergoes raising to a higher position than AspP. This is because the event variable of the VP is quantified over only after the Asp head is combined, and the VP argument does not denote an operator for *ini'* to negate; rather, *ini'* would only negate a relation. For example, the negation of a raining event in world *w* is true iff the relation does not hold. This is equivalent to saying that there is no raining event in *w*. In fact, in Mandarin, we observe that negation invariably precedes progressive aspect, as in (23) below. I assume that internal negation modifying VP or AspP produces a similar outcome.

The meaning of (12) is computed as shown in (19), which asserts that there exists a time *t'* in the coming week when there is no instance of raining events with their temporal trace included. This interpretation clearly differs from a negative future statement as in (17) or an even more concise paraphrase: ‘there is no raining event in the next week.’

- (19) $\llbracket (13) \rrbracket^{g,c} = \llbracket \text{ModP} \rrbracket^{g,c}(t^*) = \llbracket \text{musa}'_{\text{FUT}} \rrbracket^{g,c}(\llbracket \text{AspP} \rrbracket^{g,c})(t^*) = \lambda w. \exists t' [t^* < t' \ \& \ t' \subseteq \text{next week} \ \& \ \neg \exists e [\text{rain}(e)(w) \ \& \ \tau(e) \subseteq t'] \rrbracket$

In fact, the semantic infelicity here closely resembles the well-known problem raised for an existential interpretation of the English past tense (Partee 1973:602). An existential past taking scope over negation results in a trivial reading. Similarly, when the existential operator in the semantics of *musa'*_{FUT} scopes over negation, the sentence becomes trivially true, that is, (12) would be true as long as we find a time next week at which it does not rain.

In contrast, in pattern A, when the modal scoping over *ini'* encodes the interval $[t, _)$ instead of an existential quantifier, as exemplified by (9), the sentence is felicitous. As computed in (20), (9) asserts that in worlds that align with the actual world in terms of relevant circumstances and are ranked by norms, there is no event of us watering vegetables at any time following the present moment.

$$(20) \quad \llbracket (9) \rrbracket^{g,c} = \llbracket \text{ModP} \rrbracket^{g,c}(t^*) = \llbracket nway \rrbracket^{g,c}(\llbracket \text{AspP} \rrbracket^{g,c}(t^*)) = \lambda w. \exists w' [w' \in \text{MB}_{\text{Circ}}(w, t^*)] \& \neg \exists e [\text{water}(e, we, \text{vegetables})(w') \& \tau(e) \subseteq [t^*, _)]]$$

4 A parallel in Mandarin

The proposed semantic explanation for why internal negation does not scope under a future marker encoding an existential quantifier can be directly tested by Mandarin, which has an internal negator compatible only with the existence of events.

Similar to Atayal, Mandarin also employs two sentential negators, *bù* and *méi*, but their use is not solely determined by syntactic height. *Bù* is used to negate bare stative verbs, whether they are individual- or stage-level (21). In contrast, *méi* is the choice for negating bare eventives (22) as well as those marked with viewpoint aspects, e.g., *zài* 'PROG' (23) and *-guò* 'EXP'.

$$(21) \quad \begin{array}{l} \text{tā} \quad \{\text{bù}/*\text{méi}\} \quad \{\text{pàng/gāoxìng}\}. \\ \text{3SG} \quad \text{NEG} \quad \text{fat} \quad \text{happy} \\ \text{'(S)he is not \{fat/happy\}.'} \end{array}$$

$$(22) \quad \begin{array}{l} \text{zuótiān} \quad \{\text{*bù/méi}\} \quad \text{xiàiyǔ}. \\ \text{yesterday} \quad \text{NEG} \quad \text{rain} \\ \text{'It didn't rain yesterday.'} \end{array}$$

- (23) xiànzài {*bù/méi} zài xiàyǔ.
 now NEG PROG rain
 ‘It is not raining now.’

Moreover, only *bù*, but not *méi*, can negate epistemic, deontic, ability, and future modals, (24a) vs. (24b).⁷

- (24) a. **bù** yīdìng / **bù** kěnéng/ **bù** bì / **bù**
 NEG EPIST.NEC NEG EPIST.POS NEG CIRC.NEC NEG
 kěyǐ / **bù** huì/ **bù** huì
 CIRC.POS NEG ABIL NEG FUT
- b. * **méi** yīdìng / **méi** kěnéng / **méi** bì / **méi**
 NEG EPIST.NEC NEG EPIST.POS NEG CIRC.NEC NEG
 kěyǐ / **méi** huì / **méi** huì
 CIRC.POS NEG ABIL NEG FUT

Lin (2003) argues that *méi* selects eventive complements, while *bù* more strictly selects stative situations that ‘require no input of energy’. As a result of this analysis, certain statives that opt for *méi* instead of *bù* (25a), durative sentences (25b), and progressive sentences (23) would not be considered stative, whereas most modal sentences would be categorized as stative in Mandarin (24a). Lam (2022) offers an alternative perspective based on dialectal comparison and grammaticalization. According to Lam, *bù* emerged earlier than *méi*, but *méi* underwent grammaticalization and expanded its use from being an existential verb to encompass standard negation, particularly negating the existence of events. In line with Lam’s viewpoint, I assume that *bù* serves as the default negator, while *méi* is chosen over *bù* when it is internal *and* when there is an event variable in VP.

- (25) a. tā {*bù/méi} yǒu xiǎohái.
 3SG NEG have child
 ‘(S)he does not have children.
- b. qiáng shàng {*bù/méi} guà-zhe yī fú huà.
 wall above NEG hang-DUR one CLF picture
 ‘There is not a picture hanging on the wall.’ (Lin 2003:431)

⁷ The physical ability and volitional modals appear to accept either negator, {*bù*/?*méi*} *néng* ‘not physically able to’, {*bù*/?*méi*} *kěn* ‘not willing to’, possibly because of the eventive nature of these modals.

Let us now consider our prediction concerning negation within the scope of future modals. The examples (26a, b) show that the future modal *huì*_{FUT} cannot scope above either negator (in comparison to (21) and (22), respectively). The intended reading is consistently conveyed when *bù* scopes over *huì*_{FUT}.⁸ These two cases mirror the patterns C and B in Atayal, respectively (Table 1), differing only in the aspectual selection of the two negators. Notably, Atayal pattern A does not appear to exist in Mandarin, as seen in infelicitous examples like *#kěyǐ méi* ‘(intended) can not’.

- (26) a. tā yǐhòu {#huì bù / bù huì} pàng.
 3SG in.the.future FUT NEG NEG FUT fat
 ‘(S)he will not be fat in the future.’
- b. míngtiān {#huì méi / bù huì} xiàyǔ.
 tomorrow FUT NEG NEG FUT rain
 ‘It will not rain tomorrow.’

This result aligns with our expectations if *huì*_{FUT} and all circumstantial modals in Mandarin function as a future operator encoding existential quantification, resulting in a trivial reading when combined with internal negation. Furthermore, we correctly anticipate that *méi* is acceptable in future contexts where *huì*_{FUT} cannot be used, including at least in the protasis of conditionals (27a) and predictions where the evaluation time is in the future, leading to an interpretation akin to a future perfect (27b).

- (27) a. yào shì nǐ míngtiān méi lái, wǒmen jiù huì
 if be 2SG tomorrow NEG come 1PL then FUT
 gēn nǐ māma shuō.
 to 2SG mother tell
 ‘If you don’t come tomorrow, we will tell your mother.’
- b. míngnián zhè ge shíhòu, wǒ hái méi bìyè.
 next.year this CLF time 1SG still NEG graduate
 ‘I won’t have graduated by this time next year.’

⁸ A remaining issue is that with a stage-level stative verb, the *huì* > *bù* order is possible, as in (i), compared to (26a). The difference in meaning is subtle; I suspect that the modal *huì* is more grammaticalized to resemble an epistemic modal or ‘would’.

(i) tā {huì bù / bù huì} gāoxìng.
 3SG FUT NEG NEG FUT happy
 ‘(S)he will not be happy.’

5 Conclusion and typological implication

This paper argues that future-oriented modals can lexically encode either existential quantification over future times or a right unbounded interval. This is supported by their interaction with syntactically distinct types of negation. Our analysis suggests that external negation can scope over future modals with both denotations, one of which is attested in Atayal. In Mandarin, the default negator is also expected to precede and scope over future modals with an existential quantifier. In contrast, internal negation is restricted to future modals encoding an open interval, again attested in Atayal. In both languages, internal negation is not compatible with an existential future operator. Table 2 lists the proposed typology.

Table 2: Possible combinations of forward semantics and two types of negation

	External	Internal
Existential	yes (Atayal)	no (Atayal, Mandarin)
Open interval	yes (?)	yes (Atayal)

References

- Abusch, Dorit. 1998. Generalizing tense semantics for future contexts. In Susan Rothstein (ed.), *Events and Grammar*, (pp. 13–33). Dordrecht, Netherlands: Kluwer, 13–33.
- Cariani, Fabrizio & Paolo Santorio. 2018. Will done better: Selection semantics, future credence, and indeterminacy. *Mind* 127(505): 129–165.
- Chen, Sihwei. 2018. *Finding semantic building blocks: Temporal and modal interpretation in Atayal*. Doctoral dissertation, University of British Columbia.
- Condoravdi, Cleo. 2002. Temporal interpretation of modals: Modals for the present and for the past. In David I. Beaver, Stefan Kaufmann, Brady Z. Clark, & Luis D. Casillas Martínez (eds.), *The construction of meaning*. Stanford, CA: CSLI Publications, 59–88.
- Copley, Bridget. 2009. *The semantics of the future*. New York, NY: Routledge.

- Horn, Laurence R. 2001. *A natural history of negation*. Stanford, CA: CSLI Publications.
- Lam, Cherry C.-Y. 2022. Croft's Cycle in Mandarin and Cantonese throughout history and across varieties. In Ljuba Veselinova & Arja Hamari (eds.), *The negative existential cycle*. Berlin, Germany: Language Science Press, 357–401.
- Lin, Jo-wang. 2003. Aspectual selection and negation in Mandarin. *Linguistics* 41(3):425–459.
- Partee, Barbara H. 1973. Some structural analogy between tenses and pronouns in English. *The Journal of Philosophy* 70(18):601–609.
- Su, Mei-Juan. 2004. *Taiyayu Taoshan fangyan foudingci yanjiu* [Negation in Taoshan Atayal]. Master's thesis, National Hsinchu University of Education.
- Wu, Chunming. 2013. *The syntax of linking constructions in Mayrinax Atayal and Sinvaudjan Paiwan*. Doctoral dissertation, National Tsing Hua University.

HOW many topics has Hotze worked on? Echo questions, mumbling and incredulity*

EVA CSIPAK
University of British Columbia

1 Background

In her 2017 article “On the analysis of echo questions”, Marga Reis discusses a number of phenomena that she collects under the umbrella of *echo questions*. Her work focuses on German, and here I attempt to take a closer look at two of her generalizations and compare them to the facts in Dutch and English. The first is the role of discourse particles in echo questions, and the second is word-internal stress in multimorphemic wh-words.

2 Discourse particles in echo questions

Discourse particles are particles that typically express something about how the speaker’s utterance fits into the previous discourse (Eckardt 2009) or what the speaker’s attitude is towards their own utterance. As one would expect, it makes a difference whether they occur in declaratives or interrogatives. Roughly following Farkas and Bruce (2010), I will assume that declaratives are attempts to establish material in the common ground between speaker and addressee, while interrogatives both create a disjunction on the common ground and are appeals to the addressee to provide an update that eliminates some of the disjuncts.

I will use the term *echo question* to refer to a question that directly follows another person’s utterance, either for the purpose of asking for clarification or for expressing disbelief.

* As we all know, the answer to the title question is *many*. Happy birthday, dear Hotze! Thank you for unknowingly contributing to the Dutch data presented below. I also thank Sander Nederveen, who knowingly provided the rest of the Dutch data.

- (1) A: Hotze bought a yacht.
B: Hotze bought WHAT?
B': WHAT did Hotze buy?

The utterances of both B and B' count as echo questions; they only differ in the position of the wh-word. Both can be used in two distinct contexts: either in a context where B could not hear what A said, and is thus asking for clarification, or one where B did in fact hear and finds it surprising or unbelievable that Hotze bought a yacht, so asks to double check.

I assume that B is not ready to take any responsibility for the content of the echo question, so if A's utterance contains emotive content, B's echo question would leave the emotive content with A. In both A and B's utterances, the person expressing the judgment that Hotze's yacht is awesome is A.¹

- (2) A: Hotze bought an awesome yacht.
B: Hotze bought an awesome WHAT?

Some German discourse particles (DiPs) are able to occur both in declaratives and in interrogatives, such as *doch* or *wohl*.

- (3) a. Hotze hat DOCH eine Yacht gekauft.
Hotze has DiP a yacht bought
'Hotze did buy a yacht after all.'
b. Hat Hotze DOCH eine Yacht gekauft?
has Hotze DiP a yacht bought
'Did Hotze buy a yacht after all?'

As (3a) and (3b) show, stressed *doch* can occur both in declaratives and in information-seeking questions. The attitude contributed by *doch* lies with the speaker in the case of (3a) and is shared between speaker and addressee in (3b).

¹ Normally, emotive content in questions is either attributed to the speaker, or is supposed to be shared between speaker and addressee. For example, when asking *Where is Hotze wintering his awesome yacht?*, it is the speaker who finds the yacht awesome.

Other discourse particles resist occurring in one sentence type, but not the other. For example, *ja* and *halt* resist occurring in interrogatives, while *denn* resists occurring in declaratives; see below. (The exact contribution of the discourse particles to the utterance meaning is irrelevant for our purposes here — *ja* has been described as marking the content of the utterance as already known or uncontroversial, e.g. Rapp (2018), while *halt* is said to provide some kind of conclusion, see Thurmair (1991). The meaning of *denn* is notoriously difficult to describe; see Csipak and Zobel (2016) for some discussion.)

- (4) a. Hotze ist ja/ halt/ *denn in Vancouver.
 Hotze is DiP DiP DiP in Vancouver
 ‘Hotze is in Vancouver.’
 b. Ist Hotze *ja/ *halt/ denn in Vancouver?
 is Hotze DiP DiP DiP in Vancouver
 ‘Is Hotze in Vancouver?’

In her paper, Reis (2017) shows that when speaker A makes an utterance that contains a discourse particle and speaker B echoes this utterance as a question, the discourse particle remains acceptable, since it is part of the utterance that is being mirrored. We first consider an example where speaker A uses a declarative sentence containing a (declarative) discourse particle, and we observe that speaker B’s echo question can felicitously contain the same, original declarative discourse particle(s), but not the question particle *denn*.

- (5) A: Hotze hat ja/ halt/ *denn zu Tempus gearbeitet.
 Hotze has DiP DiP DiP to tense worked
 ‘Hotze has worked on tense.’
 B: Hotze hat ja/ halt/ *denn zu WAS gearbeitet?
 Hotze has DiP DiP DiP to what worked
 ‘Hotze has worked on WHAT?’

Even though B is asking a question, this question cannot contain *denn*, a particle that is acceptable in almost all information-seeking questions. The attitude contributed by the discourse particles *ja* and *halt* remain A’s attitude; they cannot be interpreted to be B’s.

Furthermore, when speaker A utters an interrogative sentence containing a question particle, speaker B's echo question can felicitously contain the same discourse particle, but not any declarative discourse particles.

- (6) A: Ob Hotze *ja/ *halt/ denn zu NPIs gearbeitet hat?
if Hotze DiP DiP DiP to NPIs worked has
'I wonder whether Hotze has worked on NPIs.'
- B: Ob Hotze *ja/ *halt/ denn zu WAS gearbeitet hat?
if Hotze DiP DiP DiP to WHAT worked has
'(You) wonder whether Hotze has worked on WHAT?'

As before, the available interpretations for B's utterance in (6) are that B either did not understand the term *NPI* in A's utterance, or B is incredulous and wants to double check. The attitude contributed by the discourse particle again remains with A in (6).

To complete the picture, let us briefly discuss a case where A did not use any discourse particles (not discussed by Reis). In this case, B cannot felicitously add any discourse particles.

- (7) A: Hotze hat zu Tempus gearbeitet.
Hotze has to tense worked
'Hotze has worked on tense.'
- B: Hotze hat *ja/ *halt/ *denn zu WAS gearbeitet?
Hotze has DiP DiP DiP to what worked
(intended) 'Hotze has worked on WHAT?'

A's (declarative) utterance contains no discourse particles, and B can add neither the question particle *denn* nor the declarative particles *ja* or *halt*. The same holds if A's preceding utterance was an interrogative, as (8) illustrates.

- (8) A: Ob Hotze zu NPIs gearbeitet hat?
if Hotze to NPIs worked has
'I wonder whether Hotze has worked on NPIs.'
- B: Ob Hotze *ja/ *halt/ *denn zu WAS gearbeitet hat?
if Hotze DiP DiP DiP to WHAT worked has
'(You) wonder whether Hotze has worked on WHAT?'

Given what we said above about the role of discourse particles, this is of course not unexpected. The original utterance and its echo belong to A, so any markers that express an attitude towards the utterance can only be licitly added by A.

We observe that the same pattern holds for Dutch: first, if A's utterance contains a discourse particle, it is possible to keep that particular discourse particle in the echo question, as in (9). Second, it is not possible to add any discourse particles that were not part of the echoed utterance, as illustrated in (10).

(9) DUTCH:

A: Hotze heeft toch/ eigenlijk/ wel/ maar een jacht gekocht.
 Hotze has DiP/ DiP/ DiP/ DiP a yacht bought
 'Hotze bought a yacht.'

B: Hotze heeft (toch/ eigenlijk/ wel/ maar) een WAT
 Hotze has (DiP/ DiP/ DiP/ DiP) a WHAT
 gekocht?
 bought
 'Hotze bought a WHAT?'

(10) DUTCH:

A: Hotze heeft een jacht gekocht.
 Hotze has a yacht bought
 'Hotze bought a yacht.'

B: Hotze heeft (*toch/ *eigenlijk/ *wel/ *maar) WAT
 Hotze has (DiP/ DiP/ DiP/ DiP) WHAT
 gekocht?
 bought
 'Hotze bought WHAT?'

Thus, Dutch works like German in this respect. Discourse particles remain with the speaker whose utterance is echoed, and thus only those particles that are part of the original utterance are acceptable.

3 Stress on wh-words with multiple syllables

In regular information-seeking questions in German, the default word stress is on the non-wh element(s) of the wh-word.

- (11) a. WaRUM hat Hotze gelacht?
why has Hotze laughed
'Why did Hotze laugh?'
b. WieVIEL hat Hotze gelacht?
how-much has Hotze laughed
'How much did Hotze laugh?'

In echo questions, this pattern is not only reversed — it is in fact *required* to stress the *wh*-element of the question word. To remind readers that these are echo questions, I have included an English sentence as a preceding utterance; this is for brevity's sake.

- (12) A: Hotze laughed *mumble* amount./ Hotze laughed [an atypically large amount].
B: Hotze hat WIEviel/*wieVIEL gelacht?
Hotze has how-much/how-much laughed
'Hotze laughed HOW much?'
(13) A: Hotze went to *mumble*. /Hotze went to a yacht sales place.
B: Hotze ist WOhin gegangen?
Hotze is where gone
'Hotze went WHERE?'

B's responses in both (12) and (13) are only acceptable if the *wh*-part of the *wh*-word is stressed. Stressing anything else is not acceptable. Reis points out that requiring stress on the *wh*-element is not tied to their in-situ position. When B utters an echo question using regular interrogative word order as in (14a) and (14b) below, the stress still needs to be on the *wh*-element.

- (14) a. WIEviel/*wieVIEL hat Hotze gelacht?
how-much/how-much has Hotze laughed
'HOW much did Hotze laugh?'
b. WOhin/*woHIN ist Hotze gegangen?
where-to/where-to is Hotze gone
'WHERE did Hotze go?'

This pattern also holds for English and Dutch. We first look at examples from English. In (15), we observe that in B's echo question, only

stress on the *wh*-word is possible. Stressing any other word is odd. Without A's preceding utterance, we can imagine a context where B wants to find out where Hotze is and would thus ask a regular, information-seeking question. In that case, a stress pattern as in B'' seems most appropriate.

- (15) A: Hotze went *mumble*.
 B: WHERE did Hotze go?
 B': *Where did HOTZE go?
 B'': *Where did Hotze GO?

(16) illustrates that for a multi-morphemic *wh*-expression, the stress needs to be on the *wh*-element and cannot be anywhere else.

- (16) A: Hotze has written *mumble* papers.
 B: HOW many papers has he written?
 B': *How MANY papers has he written?
 B'': *How many papers has he WRITten?

When we turn to Dutch, we find the same picture again. Consider the exchange in (17) — only B's utterance with stress on the *wh*-element is acceptable.

- (17) A: Hotze heeft een jacht gekocht *mumble*.
 B: WAARom heeft Hotze een jacht gekocht?
 why has Hotze a yacht bought
 'WHY did Hotze buy a yacht?'
 B': * waaROM heeft Hotze een jacht gekocht?
 why has Hotze a yacht bought
 (intended) 'Why did Hotze buy a yacht?'

Thus we have seen that English and Dutch pattern like German. Where does this pattern come from? While Reis does not propose a worked out semantic analysis of echo questions, she does propose that the stress on the *wh*-element is focus. She argues that this focus gives rise to a special kind of focus alternatives that are not normally activated. For an echo question such as *WHERE is Hotze going*, the alternatives are:

- (18) {Hotze goes where; Hotze goes there}

That is, she suggests that what is at stake is the very existence of the question *where is Hotze going* in the conversational context. Reis does not really explain what this means, so I will speculate. For echo questions with an aim to repeat information that speaker A provided but speaker B missed, by uttering the question {Hotze goes where; Hotze goes there}, speaker B is perhaps acknowledging that A and B are entertaining different context sets — while A's belief worlds contain *Hotze is there*, B's contain *Hotze is where*. In order to reach a shared Common Ground, B needs the missing information.

In the case where B did understand where Hotze is, but is asking the echo question incredulously, B might be suggesting that while technically both speakers agree that *Hotze is there*, the actual location is so unusual that it might still warrant raising the question of *Hotze is where*.

Returning to Reis' observations, one interesting consequence of this is that any wh-words which do not have a demonstrative counterpart are predicted to be bad in echo questions, since they cannot participate in forming these focus alternatives. This is indeed what Reis finds: in German, almost all wh-words have a demonstrative counterpart and can occur in echo questions. The only exception is *wieso*, which does not have a counterpart *daso* or *soso*. And indeed *wieso* cannot occur in echo questions.

- (19) A: Hotze bought a yacht because *mumble*.
B: * WIEso hat Hotze eine Yacht gekauft?
 why has Hotze a yacht bought
 (intended) 'Why did Hotze buy a yacht?'

The same is true for Dutch *hoezo* — like German, it is the only wh-word that does not have a demonstrative equivalent, and it cannot appear in echo questions.

- (20) A: Hotze bought a yacht because *mumble*.
B: * HOEzo heeft Hotze een jacht gekocht?
 why has Hotze a yacht bought
 (intended) 'Why did Hotze buy a yacht?'

Again, B cannot use *hoezo* to ask an echo question, and it is odd to try to stress the first syllable (its wh-element). Recent work by Rullmann

and Nederveen (2024) seems to support this, since they analyze *hoezo* as a metalinguistic marker. It signals that the speaker is asking about the reason for the previous speaker's utterance, not about any reasons regarding the content of that preceding utterance.

Thus, while our understanding of the semantics of echo questions is still limited, it is reassuring that three closely related languages have similar patterns with respect to how they are formed.

References

- Csipak, Eva, and Sarah Zobel. 2016. Discourse particle *denn* in the antecedent of conditionals. In *Empirical issues in syntax and semantics 11*, ed. Christopher Pinon, 31 – 60.
- Eckardt, Regine. 2009. The real, the apparent, and what is *eigentlich*. In *Structuring information in discourse: the explicit/implicit dimension*, ed. A Behrens and Cathrine Fabricius-Hansen, volume 1, 77–108. Oslo Studies in Language.
- Farkas, Donka, and Kim Bruce. 2010. On reacting to assertions and polar questions. *Journal of Semantics* 27:81 – 118.
- Rapp, Irene. 2018. Wenn man versucht, ja nichts falsches zu sagen. *Linguistische Berichte* .
- Reis, Marga. 2017. On the analysis of echo questions. *Florida Linguistics Papers* 4:1–24.
- Rullmann, Hotze, and Sander Nederveen. 2024. A *wh* discourse particle: Dutch *hoezo*. *Proceedings of the Linguistics Society of America* 2024 9(1).
- Thurmair, Maria. 1991. Zum Gebrauch der Modalpartikel “denn” in Fragesätzen. In *Linguistische Arbeiten* 260, 377–387. Tübingen: Niemeyer.

A second-last position clitic in Sm'algyax: a solution

HENRY DAVIS

University of British Columbia

COLIN BROWN

University of California, Los Angeles

1 Recapitulation

In Brown and Davis (this volume) (henceforth, B&D), we introduced the *wh*-clitic =*du* in Sm'algyax and showed the following:

- (a) The syntactic position of =*du* is high in a root clause (taking CP as its complement).
- (b) Phonologically, =*du* is an enclitic: it is phonologically integrated with a phrase to its left.
- (c) Its linear position falls into three distributional patterns, which we have characterized as *wh-placement*, *predicate placement*, and *argument placement*.

We propose here a unified explanation for these generalizations. The core of our proposal is the following:

- (1) *Morphologically, =du is a proclitic.*

Section 2 outlines how (1) can account for the three linear positions outlined in (c) (for details, see B&D), Section 3 addresses the clause-final *wh*-particle =*da*, and Section 4 concludes and outlines a number of implications of our analysis.

2 Toward an explanation

Let us consider how the idea in (1) plays out in the three attested positions where =*du* occurs. We begin with *argument placement*, wherein the *wh*-

clitic encliticizes to the transitive subject, linearizing to the left of the object (2).¹ Argument placement is schematized in (3) below.

- (2) Dzindel dmt dzapdit Meelidu
 dzindeh=l dm=t dzap-t=t Meeli=**du**=a
 IRR.when=IRR.CN PROSP=3.I make/fix-3.II=PN Mary=Q=CN
 ts'ikts'ik?
 ts'ikts'ik
 car
 'When will Mary fix the car?'

- (3) [WH V DP_A=**du** DP_O] *Argument placement*

Assuming that =*du* is base-generated at the right periphery of the root clause (B&D Section 2) we suggest that the underlying syntactic structure for (3) can be represented as in (4):

- (4) [[WH V DP_A DP_O] =**du**]

Since its phonological requirements are met here, the question immediately arises as to why =*du* cannot remain in its base-generated position. The answer is that as a *morphological proclitic*, =*du* must precede a constituent to its *right*: hence, assuming that it will choose the most local possible morphological host, it will linearize inside the (O) DP immediately to its left, as in (3). Since it will still have a phonological host to its left, its phonological requirements will also be met, and the structure will be licit.

Next, we turn to predicate placement, shown in (5) and schematized in (6):

¹ 1=first person, 3=third person, AX=agent extraction morpheme, CN=common noun connective, I=series I clitic, II=series II suffix, IRR=irrealis, PN=proper noun connective, PROSP=prospective, Q=question particle, REAS=reason subordinator, REL=relative, SG=singular, SX=subject extraction morpheme, T="Big T" verbal morpheme, TR=transitive.

- (5) Goł gan dawłdut Dzon?
 goo=ł gan dawł-t=**du**=t Dzon
 what=IRR.CN REAS leave-3.II=Q=PN John
 ‘Why did John leave?’

- (6) [WH V=**du** DP_{S/A/O}] *Predicate placement*

It should be apparent that exactly the same analysis will account for this configuration, beginning with the base structure in (7):

- (7) [[WH V DP_{S/A/O}] =**du**]

This structure is identical to (3) except that here there is only one postverbal DP rather than two. Once again, as a morphological proclitic, =*du* will be forced to linearize inside the DP to its immediate left, this time phonologically encliticizing to the verb, and deriving the correct surface form.

Now, let us turn to *wh-placement*, where =*du* attaches to the *wh*-phrase itself.

- (8) [WH=**du** [...]] *Wh-placement*

As long as the *wh*-phrase is followed only by the predicate, the account we have already given will extend naturally to *wh-placement*, since =*du* will morphologically procliticize to the predicate to its right and phonologically encliticize to the preceding *wh*-phrase, as in (9) and (10):

- | | |
|---|---|
| <p>(9) Naayu ksüüt?
 naa=du ksüü-it
 who=Q go.out-SX
 ‘Who left?’</p> | <p>(10) Naayu sibaasu?
 naa=du sibaas-u
 who=Q scare-1SG.II
 ‘Who did I scare?’</p> |
|---|---|

However, there are plenty of cases where =*du* attaches to a *wh*-phrase even when multiple constituents follow it, as for example in (11), repeated from B&D (3):

- (11) Naayu int gaba ts'ik'aaws?
 naa=**du** in=t gap[-t]=a ___ ts'ik'aaws
 who=**Q** AX=3.1 eat[-3.II]=CN split.salmon
 'Who eats split dried salmon?'

Here, we expect predicate placement; and in fact, it turns out that in such cases =*du* freely alternates between attaching to the *wh*-phrase, as in (12a) below, and attaching to the predicate, as in (12b):

- (12) a. Naadu int yoyksa nool?
 naa=**du** in=t yoyks[-t]=a nool
 who=**Q** AX=3.1 wash[-3.II]=CN dish
- b. Naał int yoyksdu nool
 naa=ł in=t yoyks-t=**du**=a nool
 who=**IRR.CN** AX=3.1 wash-3.II=**Q**=CN dish
 'Who washed the dishes?'

We suggest that the key to extending the morphological proclitic analysis to cases such as (12a) is to treat the entire string following the *wh*-phrase as *a single DP* whose internal structure is opaque to =*du*. In that case, the *wh*-clitic will be morphologically proclitic to the DP, and will phonologically encliticize to the preceding *wh*-phrase. In contrast, in cases such as (12b), the constituent following the *wh*-phrase will be CP, and =*du* will attach to the predicate, as expected.

However, in order to avoid circularity, the claim that the constituent following a *wh*-phrase + =*du* sequence is a DP rather than a CP needs to be independently motivated. Fortunately, there is a test. Recall from B&D Section 2 that *wh*-questions in Tsimshianic can either be derived by direct or indirect movement. Direct movement parallels *wh*-movement in English: the *wh*-phrase moves to a position on the left periphery of CP, leaving a clausal remnant. Indirect movement, on the other hand, involves a base-generated *wh*-predicate on the left periphery, followed by a DP argument, which usually takes the form of a headless relative clause, as exemplified in (13) and schematized in (14), repeated from B&D (7)–(8).

- (13) Godu gu yoyksis Meeli?
 goo=du [gu yoyks-i[-t]=s Meeli ___]
 what=Q REL wash-TR-3.II=PN Mary
 ‘What did Mary wash?’

- (14) [Goo [___ [___ [gu [yoyksis Meeli ___]]]]]
 [IP WH [DP pro [CP Θ_{ret} [C [IP yoyksis Meeli Θ_{ret}]]]]]

Since the argument of the *wh*-predicate in the indirect movement structure in (13) is a DP, this is exactly the configuration where we predict =*du* will encliticize to the *wh*-phrase rather than the predicate (as indeed, it does in (13)).

Furthermore, since the indirect movement structure can be readily identified by the optional presence of the *wh*-relative pronoun *gu* (B&D Section 2), we can formulate the following prediction:

- (15) *If a wh-question contains gu, =du will always attach to the wh-phrase*

This prediction is borne out. The examples in (16) show that only *wh-placement* is available for =*du* in questions containing *gu* (compare (12b) above):

- (16) a. Naadu gu int yoyksa nooŋ
 naa=**du** gu in=t yoyks[-t]=a nooŋ
 who=Q REL AX=3.I wash[-3.II]=CN dish
 ‘Who washed the dishes?’
- b. *naa=t gu in=t yoyks-t=**du**=a nooŋ
 who=IRR.CN REL AX=3.I wash-3.II=Q=CN dish

The object questions in (17) make the same point: without *gu*, =*du* can either attach to the *wh*-phrase (17a) or in penultimate position (17b), reflecting ambiguity between direct and indirect movement. With overt *gu*, however, only indirect movement is possible, and therefore =*du* must attach to the *wh*-phrase (17c); attempts to attach it to the predicate in penultimate position are ungrammatical, as shown in (17d).

- (17) a. Godu yoyksis Meeli?
 goo=**du** yoyks-i[-t]=s Meeli
 what=**Q** wash-TR-3.II=PN Mary
- b. Goł yoyksadut Meeli?
 goo=ł yoyks-i-t=**du**=t Meeli
 what=IRR.CN wash-TR-3.II=**Q**=PN Mary
- c. Godu gu yoyksis Meeli?
 goo=**du** **gu** yoyks-i[-t]=s Meeli
 what=**Q** **REL** wash-TR-3.II=PN Mary
 ‘Who washed the dishes?’
- d. *goo=ł **gu** yoyks-i-t=**du**=t Meeli
 what=IRR.CN **REL** wash-TR-3.II=**Q**=PN Mary

It still remains to be explained why the DP (relative clause) constituent in (17c) is impenetrable to =*du*, as evidenced by the ungrammaticality of (17d). Here we appeal to the notion of a phase (Chomsky 2001 and much subsequent work). One of the leading ideas behind this notion is that phases act as “chunks” for the purposes of spell-out, and once spelled out, will be opaque to further operations — in this case, to procliticization by =*du* in the morphological component. It is commonly assumed that DPs are phases, and we adopt this assumption here.

By the same token, the CP complement of =*du* is penetrable to cliticization: this means either that CP is not a phase or that =*du* is part of the same phase as its CP complement. Evidence for the latter comes from long distance extraction, where procliticization of =*du* takes place in the matrix rather than the subordinate clause, as shown in (18)–(19), repeated from B&D (40)–(41).

- (18) Ndeł małdidut Betty gooys Meeli?
 ndeh=ł mał-t-i-t=**du**=t Betty [goo-i[-t]=s Meeli ____]
 where=IRR.CN say-T-TR=**Q**=PN Betty go-TR-3.II=PN Mary
 ‘Where did Betty say Mary went?’

- (19) Goł ha'ligoodut Bettyt giindit
 goo=ł ha'ligoot-t=**du**=t Betty [=t giin-t=t
 what=IRR.CN think-3.II=Q=PN Betty =3.I give-3.II=PN
 Michaelt Henry?
 Michael=t Henry ___]
 Michael=PN Henry
 ‘What does Betty think Michael gave Henry?’

In long-distance questions, =*du* always appears in the matrix rather than an embedded CP (see B&D Section 2.2). The inaccessibility of subordinate CPs follows if, like DPs, they constitute phases. At the point of clitic linearization, embedded CPs have already been spelled out, whereas the matrix CP has not, and is therefore accessible to =*du* placement.

Finally, as we saw in B&D Section 4, linearization of =*du* only pays attention to the predicate and its DP arguments. If we treat =*du* as a phrasal proclitic, as seems necessary to account for its positioning with respect to DPs, this distribution appears odd, since the predicate is a head. However, the only cases where =*du* apparently procliticizes to a predicate are precisely those where we have just shown that the *wh*-phrase is itself a predicate taking a DP argument (i.e., cases of indirect movement, such as in (16a) and (17c) above). In other words, here =*du* is a morphological proclitic to DP, just as in cases of argument placement. This means that we can now eliminate the predicate from the set of possible proclitic hosts, leaving us with a starkly simple generalization:

- (20) *Only DPs count for the linearization of =du*

We take it as a virtue of the current analysis that what appears at first sight to be a very complex distributional pattern is reducible to the interaction of simple constraints on linearization such as (20), together with independently motivated structural properties of the language (the distinction between direct and indirect \bar{A} -movement) and widely accepted conditions on the interface (spell-out by phase).

3 =*Da*

We have seen how the dual status of =*du* as a phonological enclitic and a morphological proclitic accounts for its “second last position” behaviour:

it will always end up sandwiched between a phonological host to its left and a morphological host (a DP) to its right. But what happens when there is only a single constituent to attach to?

It turns out that in these cases, the *wh*-question clitic takes a separate form, =*da*, which is uniformly *enclitic*.² This form is obligatory in reduced questions consisting of just a *wh*-phrase:

- | | | | |
|------|--|------|--|
| (21) | Naaya?/Naada?
naa= da
who= Q
'Who (is it)?' | (22) | Goya?/Goda?
goo= da
what= Q
'What (is it)?' |
|------|--|------|--|

=*Da* is also optional instead of =*du* in some non-reduced questions, where it always surfaces in final position.³

- | | | | |
|------|--|------|--|
| (23) | Goł gabida?
goo=ł gap-i-t= da
what=IRR.CN eat-TR-3.II= Q
'What did s/he eat?' | (24) | Naal int
naa=ł in=t
who=IRR.CN AX=3.I
gapda?
gap-t= da
eat-3.II= Q
'Who ate it?' |
|------|--|------|--|

The distribution of =*da* provides further indirect support for our analysis of =*du*, since it surfaces exactly where we expect =*du* to be impossible.⁴

² The morphophonology of =*da* is consistent with that of =*du* as described in B&D §3 (for example, =*da* also optionally exhibits free variation between [da] and [ja] when immediately following a *wh*-word), suggesting that the two are allomorphs. For reasons of space, we do not explore this possibility further here.

³ Final position is typical for question clitics across Tsimshianic: in Sgüüxs (Southern Tsimshian), both *wh*-Qs and yes-no Qs are marked by a final enclitic =*i*, and in Interior Tsimshianic, yes-no Qs are marked by a final =*aa*, with no marking for *wh*-Qs.

⁴ There is more to say about the distribution of =*da* in cases where =*du* is also possible. To be specific, =*da* is available as a (preferred) alternative to =*du* in cases of A and O extraction with a third person and no following DP, such as

4 Conclusion and Further Implications

The main points of the analysis are summarized in (i)–(vi) below.

- (i) =*du* is base-generated in the syntax in a high MoodP on the right periphery which encodes illocutionary force and takes a root CP as its complement (B&D Section 2)
- (ii) =*du* is a phonological enclitic which attaches to a prosodic host to its left (B&D Section 3)
- (iii) =*du* is a morphological proclitic which must precede a phrasal host to its right (B&D Section 4, Section 2 of this paper)
- (iv) as phases, DP and CP are opaque to cliticization once spelled out (Section 2 of this paper)
- (v) only a DP may serve as a proclitic host for =*du* (Section 2 of this paper)
- (vi) where the dual requirements of =*du* as a morphological proclitic and a phonological enclitic cannot be met, =*da* (which is both a morphological and a phonological enclitic) is inserted instead (Section 3 of this paper)

In this final section, we briefly explore some of the implications of our account and the model of grammar which it entails. While for reasons of space we cannot give an explicit formal analysis, we will point to the kind of grammatical architecture which we think will be necessary to handle the Sm'algyax facts.

We begin with a significant theoretical claim which we think is almost unavoidable, given the facts we have presented.

- (25) *The linearization of clitics is not reducible to either their syntax, their phonology, or any combination of the two.*

those in (23) and (24); otherwise, only =*du* is permitted. We set aside this extra complication here for the sake of space.

Indeed, insofar as our account is successful, it serves as a proof-of-concept of the existence of a separate morphological component of the grammar responsible for the linearization of clitics.

Second, we observe that this component must occupy a very specific position in the architecture of the grammar: it takes the syntax as its input, and the phonology as its output. Clitic linearization, in other words, takes place at the interface between syntax and phonology.

Third, our account supports a two-step model of lexicalization, in which features governing linearization are activated at the first step, and morphophonological features (e.g., those governing contextual allomorphy) come into play in the phonological component only after linearization has taken place.⁵

Fourth, we have outlined some of the parameters of the linearization operation itself. The following factors appear to be relevant:

- (a) The direction of (morphological) cliticization (left for enclitics, right for proclitics).⁶
- (b) The nature of the (morphological) host. There are two factors to consider here:
 - (i) Whether the host is a phrase (leading to “phrasal affixation”) or a head.
 - (ii) The categorial signature of the host.

For =*du*, the direction is rightward, and the host is a DP.

In its appeal to morphology, the model we have briefly outlined here owes an intellectual debt to previous accounts of cliticization such as those of Billings (2002), Anderson (2005), and particularly Klavans (1985). In fact, a significant empirical contribution of this paper is to vindicate one of the predictions made by Klavans’ parametrization of possible clitic

⁵ A model of clitic linearization with exactly these properties is laid out in Huijsmans (2023) on the basis of evidence from second position clitics in Salish: see also Davis and Huijsmans (2021).

⁶ Given the existence of what Mulder and Sellers (2010) refer to as “flexiclitics” in Sm’algyax (that is, clitics which indiscriminately attach either to the left or right), it is possible that this parameter can remain unspecified for some clitics.

positions: the existence of penultimate or “second-last” position clitics (Type 5 in her typology: see Klavans 1985: 103).⁷

However, the architecture which we employ and the division of labour between the narrow syntax and its interfaces is very much in the derivational tradition of minimalism, as is our use of the machinery of spell-out by phase. Overall, we hope to have shown here that a separate operation of clitic linearization in the morphology allows for an elegant account of a very complex pattern of cliticization in Sm'algyax, with broader implications for the treatment of clitics cross-linguistically and the architecture of the syntax-phonology interface.

References

- Anderson, Stephen. 2005. *Aspects of the theory of clitics*. Oxford University Press.
- Billings, Loren A. 2002. Phrasal clitics. *Journal of Slavic Linguistics* 10:53–104.
- Brown, Colin, and Henry Davis. This volume. A second-last position clitic in Sm'algyax: a puzzle.
- Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A life in language*, 1–52. MIT Press.
- Davis, Henry, and Marianne Huijsmans. 2021. Clitics and coordination in two Salish languages. Poster presented at WCCFL 39, University of Arizona.
- Huijsmans, Marianne. 2023. Second-position clitics, from morphosyntax to semantics: The ʔayʔajuθəm (Comox-Sliammon) perspective. Doctoral Dissertation, University of British Columbia.
- Klavans, Judith L. 1985. The independence of syntax and phonology in cliticization. *Language* 61:95–120.
- Mulder, Jean, and Holly Sellers. 2010. Classifying clitics in Sm'algyax: Approaching theory from the field. In *Fieldwork and linguistic documentation in indigenous languages of the Americas*, ed. An-

⁷ See Billings (2002) for criticism of Klavan's original evidence for this claim, based on data from the Australian indigenous language Nganhcara.

drea L. Berez, Jean Mulder, and Daisy Rosenblum, 2, 33–56. URL <http://nflrc.hawaii.edu/lc/>, Language Documentation and Conservation Special Publication.

A pilot study: Affect and grammatical anomaly in discourse*

VEENA DWIVEDI
Brock University

Abstract: We investigated whether differences in emotional temperament ('dispositional affect') would influence question-response accuracy rates for sentences containing modal auxiliaries. Modal sentences were embedded in contexts that were either hypothetical or factual (control). Modal auxiliaries either required clauses to restrict their interpretation ('dependent modals', *might/would*) or did not ('independent modals' *must/should*) (Stump 1985). 49 participants read two-sentence discourses followed by superficial true/false descriptive statements, e.g., *The art collector is admiring the statue. It would cost thousands of dollars.* Statement: *The art collector is appreciating the statue.* 1) True 2) False. We replicated previous work showing a cost for dependent modals when embedded in factual vs. hypothetical contexts. Moreover, low positive affect individuals were more accurate responding to independent vs. dependent modal sentences, regardless of context type. We interpret these findings as a facilitation effect for low positive affect individuals, who prefer simpler structures for task requirements.

Keywords: dispositional affect, modal auxiliaries, PANAS, sentence processing, question response accuracy

1 Introduction

In recent work, we have shown that the language comprehension system is not independent of the affective system in the mind/brain. Recent Event-Related Potential (ERP) work in our lab (Selvanayagam et al. 2020; Dwivedi 2020; Dwivedi & Selvanayagam 2021; Selvanayagam et al. 2019) has shown that neural responses to sentences differed according to dispositional affect scores as measured by the Positive and Negative Affective Schedule ("PANAS", Watson et al. 1988). Other ERP language studies have also found correlations between emotional mood (via mood induction procedures) and ERP components to sentences (see Chwilla et al. 2011; Federmeier et al. 2001; Vissers et al. 2010, 2013).

* Dear Hotze: We are a long, long way from sitting in class at UMass (not to mention our semantics circle club!). You heard the earliest versions of my ideas on modal subordination sentence processing during seminar – and at my kitchen table. Thanks for your comments back then and hope you enjoy the 21st century version! Happy Birthday. Contact info: vdwivedi@brocku.ca

These language specific findings corroborate work found in other cognitive domains (e.g., attention, MacLean et al. 2010; visual cognition: Schmitz et al. 2009), as well as cognitive processing in general (Huntsinger & Isbell 2014) regarding the role of affect.

Work from our lab suggests, at first blush, that the positive affective system is associated with the structural component of language. In Selvanayagam et al. (2019), we conducted an ERP dual-task study examining quantifier scope ambiguous sentences. Participants had to press either '1' or '2' at the presentation of a word in blue font on the computer screen, to indicate whether one or two words were presented. Sentence types were of the form *Every/The kid climbed a/the tree/trees*. The original hypothesis (Dwivedi & Gibson 2017; see Patson & Warren 2010) was that sentence interpretation effects would interfere with task requirements. That is, we expected interference effects when the (plural) word *trees* required a '1' button press when it happened to be the only word on the screen. At first, we wanted to know whether this potential difficulty would be mirrored at *tree* when it might be interpreted as covertly plural in quantifier scope ambiguous sentences (see Dwivedi & Gibson for discussion of results). When this work was followed up with an investigation with dispositional affect, we found P300 effects at *tree(s)* for all button-press conditions; this ERP component is known to be elicited in dual task studies. Interestingly, we observed that the sentence with the least amount of information relevant to the task, *The kid climbed the tree*, showed smaller ERP responses for low positive affect individuals (resulting in a larger P300 effect overall). We described this negative correlation between positive affect and P300 ERP amplitude differences in terms of individuals' motivation for sentence meaning interpretation. We speculated that rather than deeply attend to the meaning of the sentences, low positive affect individuals were primarily concerned with task requirements (and grammatical information relevant for the task).

In other work (Selvanayagam et al. 2020; Dwivedi 2020), we found high positive affect individuals showed larger P600 effects (Hagoort et al. 1993; Osterhout & Holcomb 1992). This ERP component was found in response to classic reduced relative clause sentences such as *The broker planned/persuaded *to conceal the transaction *was sent to jail*; (frontal) P600 effects were elicited at *was*. In that experiment, every critical sentence was followed by comprehension questions, such as: *Was the broker concealed/persuaded?* 1) Yes 2) No. Larger P600 effects were found for high positive vs. low positive individuals. We hypothesized that high positive affect individuals would be more motivated to revise

sentences that exhibited errors, whereas low positive affect individuals would not. This fits well with the theoretical notion of the P600 component as an index of syntactic revision (Kaan & Swaab 2003).

Together these results suggest that the positive affect system is sensitive to information that is grammatically relevant for task requirements.

We decided to explore this idea further by investigating another grammatical effect in language, now involving modal auxiliaries in discourse. In previous work, we examined grammatical requirements across discourse (as dictated by the semantic component) and observed an empirical contrast between sentences with *would* vs. *should* modal auxiliaries. That is, whereas modal auxiliaries such as *would* require a non-factual restrictor to be interpreted, modals such as *should* do not. This grammatical contrast was observed empirically in a self-paced reading study (Dwivedi 1996). That is, increased reading times were associated with sentences containing *would* that were preceded by factual (control) context sentences which were incongruent with the modal's requirements, e.g., *My friend's business will hire a new salesperson. The position would be open in May.* In contrast, ease of processing was observed when the previous context sentence was hypothetical (and therefore congruent with grammatical expectations), as in: *Maybe my friend's business will hire a new salesperson. The position would be open in May.* This contrast regarding different context types (control vs. hypothetical) was not observed for *should* sentences, e.g., *Kevin will try to find a date for the party. He should try a dating service* vs. *Perhaps Kevin will try to find a date for the party. He should try a dating service.* In that work, the contrast between *would* and *should* sentences was attributed to the idea that *would* requires an "if-clause" type of an antecedent to be interpreted (Stump 1985). That is, the meaning of the previous *would* discourse is something like, *Maybe my friend's business will hire a new salesperson, and [if that is the case, then] the position would be open in May* (Roberts 1996). No such 'if-clause' type of restrictor is necessary for the interpretation of sentences containing *should*. We later followed up on this work using ERP methods (Dwivedi et al. 2010; Dwivedi et al. 2006). Interestingly, when the grammatical requirements were not met in control contexts, a 'semantic' P600 effect emerged. We argued that this ERP effect, typically associated with morpho-syntactic anomaly and/or garden-path sentence types (Hagoort et al. 1993; Osterhout & Holcomb 1992), was indexing a grammatical requirement not being met during interpretation. This was among the first

papers to show, using ERPs, that structural effects could be observed using formal semantic constructs.

In the present work, we followed up on these previous modal auxiliary experiments by examining whether a similar grammatical contrast would be observed when we expanded the modal types to also include *might* vs. *must*. That is, like epistemic *would*, the possibility modal *might* also requires a non-factual restrictor for interpretation, in contrast to *should* and *must* (Stump 1985). Thus, we examined question-response accuracy rates after two-sentence discourses, where the context sentence was either factual (control) or hypothetical, and the continuation sentence contained one of two modal types: modals that were dependent on context for interpretation, ‘dependent’ *might*, *would* vs. ‘independent’ modals that were not, *must*, *should*.

We tested two sets of hypotheses. First, we hypothesized that we would replicate previous findings, such that empirical contrasts would be observed for dependent modal sentences embedded in control vs. hypothetical contexts. No such contrast was expected for independent modal sentences. That is, we expected higher accuracy rates when dependent modal sentences were embedded in contexts that were congruent with grammatical expectations, where no such difference for independent modals was expected. Next, regarding affect: given that we have indicated that dependent vs. independent modals have different requirements dictated by the grammatical component, we expected that positive affect scores should correlate with question-response accuracy rates. It could be the case, following our P600 results with reduced relatives (Dwivedi 2020), that high positive affect individuals are more sensitive to grammatical contrasts found for dependent modals. If so, a negative correlation is expected for positive affect and dependent modal sentences, where high positive affect individuals would perform more poorly when grammatical expectations were unmet in control contexts. Thus, these individuals were expected to show lower accuracy rates for *might*, *would* conditions when embedded in control vs. hypothetical contexts. On the other hand, based on our quantifier scope study (Selvanayagam et al. 2019), it could be the case that a negative correlation would be found for independent modals, when these were embedded in factual (control) contexts. That is, low positive affect individuals would be more accurate at *must*, *should* conditions when these are embedded in control contexts, since these discourses would have the least amount of grammatical structure and information. As such, the form of these discourses would be congruent with low positive affect

individuals' processing preferences, resulting in higher question response accuracy rates.

2 Materials and methods

2.1 Ethics statement

This study received ethics approval from the Brock University Social Science Research Ethics Board (SREB) prior to the commencement of the experiment (REB 16-179). Written, informed consent was received from all participants prior to their participation in the experiment.

2.2 Participants

Forty-nine right-handed native speakers of English (45 female, mean age 19.0 years, range 18 to 25 years). were recruited via the Brock University SONA participant pool and posters; participants were given partial course credit or were paid \$10 (if not eligible for course credit).

Table 1: Examples of different modal stimuli conditions with true/false questions

		Modal Type	
		Independent	Dependent
Hypo- tical context	S1: For all we know, the forester is looking for a hibernating bear. S2: It should rise after the snow melts.	S1: The advertiser is conceiving of a possible campaign. S2: It would turn around the company	
	<i>It's possible that the forester is looking for an old growth forest.</i> 1) True 2) False	<i>The advertiser is thinking about a career change.</i> 1) True 2) False	
Control context	S1: The technician is installing an antenna. S2: It must supply a clear signal.	S1: The firemen are examining the ladder. S2: It might rise from the back of the trunk.	
	<i>The technician is erecting an antenna.</i> 1) False 2) True	<i>The firemen are inspecting the ladder.</i> 1) True 2) False	

2.3 Materials

Each experimental trial consisted of two sentences followed by a statement requiring a true/false response. The first sentence was the context sentence (Sentence 1, S1), which was either hypothetical (i.e., non-factual) or control (i.e., factual). This was followed by a continuation sentence (Sentence 2, S2), which contained one of four modal auxiliaries: those requiring restrictive clauses for interpretation (*might, would*) vs. those that do not (*must, should*). The former modal auxiliaries are dependent on context for interpretation, whereas the latter are not (independent). Sentences were adapted from (Dwivedi et al. 2006). Thus, the factorial combination of context type (control vs. hypothetical) and continuation sentence type (independent vs. dependent modal) yielded four conditions, see Table 1. There were 16 items in each condition, where half of each condition used either *might/would* or *must/should*. As this was a pilot study, stimuli length was not controlled for, and each cell had a different sentence type.

Hypothetical context sentences differed from control contexts in that they contained markers of non-factual mood (such as a modal adverb *possibly, likely, perhaps*, etc. and/or a non-factive propositional attitude verb such as *consider, muse, wonder*, etc.). In addition, the context sentence also used a verb of creation (such as *paint, bake, write*) to further bias for a non-specific reading of the indefinite noun phrase (NP) object. The control (factual) context sentences did not contain modal adverbs or non-factive propositional attitude verbs and used verbs of using (such as *read, show, enjoy*).

All 64 stimuli were followed by a statement requiring a True/False response; there were an equal number of True/False responses across trials and the position of True/False on the screen was also counterbalanced.

Four lists were created to ensure that the conditions were counterbalanced as per Latin square design. The 64 experimental items were combined with 24 stimuli from an unrelated experiment (see Dwivedi 2013), and 100 fillers, for a total of 188 items per list. All stimuli were followed by forced choice questions or true/false statements. Two buttons (labeled as “1” and “2”) were designated for answer selection. An example filler stimulus/question pair is shown in below:

- (1) S1: Because of the thunderstorm, Lara had trouble sleeping.
S2: She felt terrible the next day.
Q: Did Lara sleep well?

1) Yes 2) No

Participants pressed the button that corresponded to the answer on the screen. Answers were counterbalanced such that equal numbers of correct answers were displayed on the right and left side of the screen.

2.4 Procedure

Upon arrival for the experimental session, participants were given three short written questionnaires to complete (in counterbalanced order) regarding (i) reading habits, (ii) a handedness inventory (Briggs & Nebes 1975), and (iii) the PANAS (Watson et al. 1988) before the start of the self-paced reading study. Before starting the experiment, participants practiced on a short list of items to familiarize themselves with task requirements. The study used a moving window display (Just et al. 1982), presented via E-prime software. Questions were presented in their entirety with potential answers on the same screen, after participants has read the critical sentence. Participants controlled the timing of the presentation of the question, and upon answering the question, the next stimulus appeared after 1200 milliseconds.

The order of sentence presentation was randomized per participant by E-Prime software. A 19" widescreen Dell LCD monitor was approximately 18–24 inches from the participant, level with the participant's point of view.

Participant responses were recorded via a PSTnet serial response button box. The experiment lasted approximately 30 minutes, and participants were debriefed after the session as to the nature of the experiment.

3 Results

Given that this was a pilot study, length of sentences was not controlled for. As such, measures collected for sentence reading times are not of interest here and will not be described.

We focus on question-answer responses only.

3.1 Filler comprehension questions

Comprehension rates for questions at filler conditions were at ceiling, 96.54% ($SD = 3.52\%$), and contrasted with the overall accuracy rate for stimuli with modals 92.92% ($SD = 3.80\%$). A paired samples *t*-test

revealed significant difference between these accuracy rates, indicating a level of difficulty with sentences with modal auxiliaries; $t(48) = -6.04$, $p < .001$, $d = -0.86$.

3.2 Experimental trial comprehension questions

Results for accuracy rates (%) for independent modals (*must, should*) revealed that control contexts ($M = 91.96\%$, $SD = 5.71\%$) were responded to less accurately than hypothetical contexts ($M = 94.13\%$, $SD = 6.17\%$), although this did not reach significance, $t(48) = -1.97$, $p = .055$, $d = -.28$. In contrast, dependent modals (*might, would*) revealed a strong difference (as indicated via Cohen’s d), where control contexts ($M = 90.82\%$, $SD = 7.77\%$) were responded to at a significantly lower rate vs. hypothetical contexts ($M = 94.77\%$, $SD = 5.16\%$), $t(48) = -3.40$, $p < .001$, $d = -.49$.

3.3 Correlational analyses

Positive Affect (PA) scores ranged from 17 to 41 ($M = 30.7$, $SD = 5.7$); Negative Affect (NA) scores ranged from 11 to 43 ($M = 19.6$, $SD = 6.1$).¹ Table 2 shows Pearson r correlations with Positive Affect scores and question-response accuracy rates in each condition

Table 2: Pearson correlations for question-response accuracy rates between PA vs. hypothetical/control and independent/dependent modals

		Control: Independent	Hypothetical: Independent	Hypothetical: Dependent	Control: Dependent
PA	r	-.36*	-.31*	-.20	-.04
	p	.012	.032	.162	.778

Note. * $p < 0.05$

¹ For the sake of completeness, we did also run correlations between question-response accuracy rates and Negative Affect (NA) scores (range from 11 to 43; $M = 19.6$, $SD = 6.1$). No significant correlations were observed with NA. Factual-independent returned a correlation of $-.05$ ($p = .722$), factual-dependent had a correlation of $-.22$ ($p = .132$), hypothetical-independent had a correlation of $.21$ ($p = .149$), and hypothetical-dependent had a correlation of $-.12$ ($p = .404$).

Interestingly, moderate negative correlations were observed between PA scores and independent modals — regardless of context type. Nothing was found for dependent modal conditions. Given this result, we ran a correlational analysis for modal type and PA. The overall mean accuracy for independent modals was 93.21% ($SD = 4.45\%$), vs. dependent modals where the value was 92.97% ($SD = 5.12\%$). Figure 1 shows a relatively strong negative correlation between PA scores and question-response accuracy for independent modals ($r(47) = -0.43, p = .002$) (likely due to increased power due to increased number of items). Thus, participants with smaller PA scores had higher accuracy rates (conversely, participants with larger PA scores had lower accuracy rates) for independent modals, where no relationship was found for dependent modals. As expected, no correlation between dependent modals and PA scores was found ($r(47) = -0.13, p = .371$) (see Figure 2).

Figure 1: Correlation between question-response accuracy for independent modals and PA

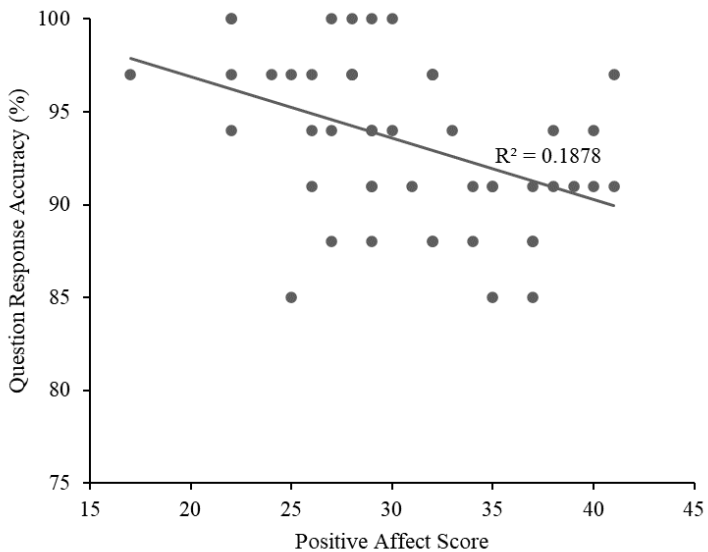
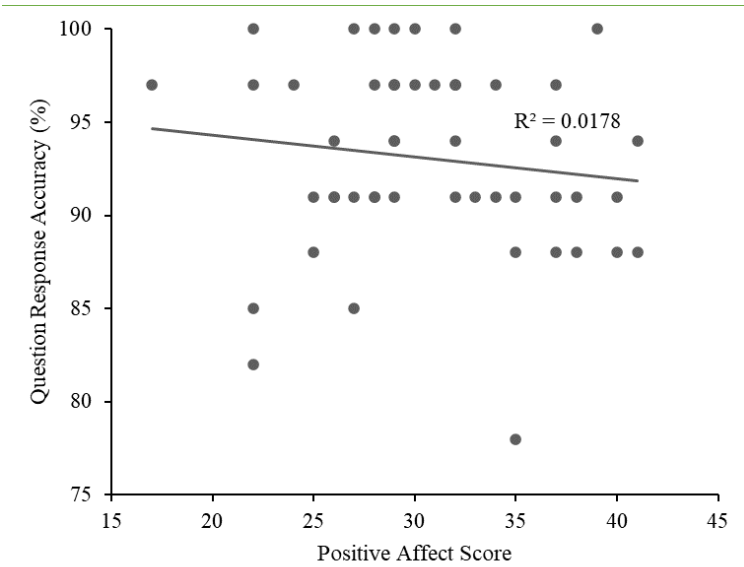


Figure 2: Correlation between question-response accuracy for dependent modals and PA



4 Discussion

We note that overall, the results showed that question-response accuracy rates were numerically higher in hypothetical vs. control contexts. As predicted, a robust (as indicated via effect size) difference for question-response accuracy rates was observed for dependent (*might, would*) vs. independent (*must, should*) modal sentences. The former modal type is preferably interpreted with non-factual restrictors whereas the latter need not have one for interpretive purposes. As such, when the discourse structure is incongruent with grammatical expectations, there is a cost — question-response accuracy rates were lower for dependent modal sentences embedded in control contexts. These results confirm the self-paced reading time findings of Dwivedi (1996), as well as the ERP findings of Dwivedi et al. (2006) and Dwivedi et al. (2010). This finding on its own is of interest for a few reasons. First, we note that the participants in the original study (Dwivedi 1996) were American English speakers in Massachusetts, circa 30 years ago, in contrast to more recent Canadian participants in Ontario. It is interesting to note that this grammatical contrast has not changed in time or via geographical considerations. Next, in an era of handwringing regarding the replication

crisis in psychology (see Schmidt & Oh 2016, among others), it is heartening to see an effect replicated across multiple methods and measures, over a span of several decades.

4.1 Low positive affect and structure

The present finding that low positive affect individuals respond differently to questions after independent vs. dependent modals supports our claims regarding findings in Selvanayagam et al. (2019). There, we proposed that individuals with low positive affect are not as engaged with sentence interpretation, and instead are focused on task accuracy. That is, rather than focusing on interpreting the sentences, they are focused on completing task requirements with as little effort as possible. Independent modals (*must, should*) do not require restrictive clauses for interpretive purposes (i.e., these are grammatically simpler, and/or have grammatically simpler discourse structures). As such, when sentences containing independent modals are presented (where these, by definition require less structure for interpretation), a facilitation effect emerges for low positive affect individuals — resulting in better question-response selection (Szucs & Soltész 2007). That is, the form and interpretation of the sentence stimuli allow for a more accurate response on behalf of low positive affect individuals since the stimuli are congruent with participant preferences for cognitive processing. This proposal would help explain the higher accuracy rates for independent modals for low positive affect individuals (i.e., negative correlation), where no relation is found for dependent modals. Because independent modals require less grammatical structure for interpretation, these are preferred by low positive affect participants. A carefully controlled follow-up study should be conducted to confirm this finding.

5 Concluding remarks

In sum, we investigated question-response accuracy rates to dependent (*might, would*) vs. independent (*must, should*) modal auxiliary sentences embedded in hypothetical vs. control contexts. We investigated two hypotheses: first, whether we would replicate previous findings regarding ease of processing when dependent modals were embedded in hypothetical contexts vs. control (factual) contexts. Second, we wanted to know whether positive affect would correlate with question-response accuracy rates. We did replicate our previous work showing a cost to interpretation when dependent modals were embedded in control

(factual) vs. hypothetical contexts. Our results also showed a negative correlation between question-response accuracy rates and positive affect for independent vs. dependent modals. We interpreted these findings as a facilitation effect — sentence stimuli that had fewer grammatical (and therefore structural) requirements for interpretation would be preferred for participants whose main focus was on task accuracy vs. sentence interpretation. These preliminary findings are among the first to relate dispositional affect to individual differences in sentence interpretation.

References

- Briggs, Gary G. & Robert D. Nebes. 1975. Patterns of hand preference in a student population. *Cortex* 11:230–238.
- Chwilla, Dorothee J., Daniele Virgillito, & Constance T. W. M. Vissers. 2011. The relationship of language and emotion: N400 support for an embodied view of language comprehension. *Journal of Cognitive Neuroscience* 23(9):2400–2414.
- Dwivedi, Veena D. 2020. Affect and syntactic anomaly. Poster presentation for the 42nd annual meeting of the *Cognitive Science Society*.
- Dwivedi, Veena D. & Janahan Selvanayagam. 2021. Effects of Dispositional Affect on the N400: Language Processing and Socially Situated Context. *Frontiers in Psychology* 12. <https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2021.566894>
- Dwivedi, Veena D. 1996. Modality and discourse processing. In Silvina Montrul & Martine Kessler Robb (eds.), *McGill Working Papers in Linguistics Vol. 12*. Montreal, QC: Department of Linguistics, McGill University, 17–52.
- Dwivedi, Veena D, John E. Drury, Monika Molnar, Natalie A. Phillips, Shari Baum, & Karsten Steinhauer. 2010. ERPs reveal sensitivity to hypothetical contexts in spoken discourse. *NeuroReport* 21(11): 791–795.
- Dwivedi, Veena D., Natalie A. Phillips, Maude Laguë-Beauvais, & Shari Baum. 2006. An electrophysiological study of mood, modal context, and anaphora. *Brain Research* 1117(1):135–153.
- Federmeier, Kara D., Donald A. Kirson, Eva M. Moreno, & Marta Kutas. 2001. Effects of transient, mild mood states on semantic memory

- organization and use: an event-related potential investigation in humans. *Neuroscience Letters* 305(3):149–152.
- Hagoort, P., Colin M. Brown, & Jolanda Groothusen. 1993. The syntactic positive shift as an ERP measure of syntactic processing. *Language and Cognitive Processes* 8:439–483.
- Huntsinger, Jeffrey R., & Linda M. Isbell. 2014. The Affective Control of Thought: Malleable, Not Fixed. *Psychological Review* 121(4):1–75.
- Kaan, Edith & Tamara Y. Swaab. 2003. Electrophysiological evidence for serial sentence processing: A comparison between non-preferred and ungrammatical continuations. *Cognitive Brain Research* 17(3), 621–635.
- MacLean, Mary H., Karen M. Arnell, & Michael A. Busseri. 2010. Dispositional affect predicts temporal attention costs in the attentional blink paradigm. *Cognition and Emotion* 24(8):1431–1438.
- Osterhout, Lee & Phillip J. Holcomb. 1992. Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language* 31(6):785–806.
- Schmidt, Frank L. & In-Sue Oh. 2016. The Crisis of Confidence in Research Findings in Psychology: Is Lack of Replication the Real Problem? Or Is It Something Else? *Archives of Scientific Psychology* 4(1):32–37.
- Schmitz, Taylor W., Eve De Rosa, & Adam K. Anderson. 2009. Opposing Influences of Affective State Valence on Visual Cortical Encoding. *Journal of Neuroscience* 29(22):7199–7207.
- Selvanayagam, Janahan, Brent Dryczewycz, Louis Schmidt, Veena Dwivedi. 2020. P600 and dispositional affect. Poster presentation for the 27th annual meeting of the Cognitive Neuroscience Society.
- Selvanayagam, Janahan, Victoria Witte, Louis A. Schmidt, & Veena D. Dwivedi. 2019. A preliminary investigation of dispositional affect, the P300, and sentence processing. *Brain Research* 1721.
- Stump, Gregory T. 1985. *The semantic variability of absolute constructions*. Dordrecht, Netherlands: D. Reidel Publishing Company.

- Szucs, Dénes & Fruzsina Soltész. 2007. Event-related potentials dissociate facilitation and interference effects in the numerical Stroop paradigm. *Neuropsychologia* 45(14):3190–3202.
- Vissers, Constanc. T. W. M., Uli G. Chwilla, Jos I. M. Egger, & Dorothee J. Chwilla. 2013. The interplay between mood and language comprehension: Evidence from P600 to semantic reversal anomalies. *Neuropsychologia* 51(6):1027–1039.
- Vissers, Constance. T. W. M., Daniele Virgillito, Daniel A. Fitzgerald, Anne E. M. Speckens, Indira Tendolkar, Iris van Oostrom, & Dorothee J. Chwilla. 2010. The influence of mood on the processing of syntactic anomalies: Evidence from P600. *Neuropsychologia* 48(12):3521–3531.
- Watson, David, Lee A. Clark, & Auke Tellegen. 1988. Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology* 54(6):1063–1070.

Enough!

KAI VON FINTEL

Massachusetts Institute of Technology

1 Introduction

Beck and Rullmann (1999) argued that a simple maximality based account to questions is insufficient (pun intended) for a question like (1):

- (1) How many eggs are sufficient (to bake this cake)?

This question does not ask for the maximal number of eggs that is sufficient for baking this cake. Instead, Beck & Rullmann proposed a more sophisticated maximal informativity account, according to which (1) asks for the most informative number n such that n eggs are sufficient for the cake. This will in fact be the minimum number eggs needed.¹

Along the way, Beck & Rullmann discussed the notion of sufficiency, proposing ideas that had not been made explicit before. They did this not because sufficiency is a primary target of their investigation but to make sure that the technical implementation of their theory of maximal informativity of questions is explicit and plausible.

They suggested two equivalent paraphrases of the sentence in (2):

- (2) Four eggs are sufficient (to bake this cake).
- (3) a. It is not necessary (given the rules for your cake baking) that you have more than four eggs.
- b. It is possible (given the rules for your cake baking) that you have only four eggs.

When they specified the lexical meaning of *sufficient*, they opted for using the \Diamond only version:

¹ The maximal informativity account is extended in von Fintel, Fox, and Iatridou (2014) to the semantics of definites.

We will derive this semantics via the lexical meaning of *sufficient*. We will take as our guideline the paraphrase in (3b). We will assume that semantically the argument of *sufficient* is propositional in nature. *Sufficient* then contributes modal possibility as well as a meaning component amounting to *only*.

Their brief discussion was pivotal for my work with Sabine Iatridou on the sufficiency modal construction (SMC; von Fintel and Iatridou (2005a, 2007)), where we looked at ways that language compositionally constructs a sufficiency meaning. The central construction we analyzed displays a structure that seems to correspond to neither (3a) nor (3b):

(4) To get good cheese, you only have to go to the North End!

We argued that (4) actually does correspond to the structure of (3a): we proposed that in (4), *only* decomposes into negation + *more than* and wrap around the necessity modal *have to*. We identified a bunch of properties of the SMC and noted many puzzles and connections.²

In the intervening years, there has been quite a bit of work on sufficiency (and at least one of its foils: excess) but I believe that the domain is ripe for further work. Consider this then an invitation for Hotze to (re-)join the fray and clarify things for us all.³

2 The notional category of sufficiency

The study of the notional category of sufficiency, or “enoughology”, promises to provide a field of inquiry unsurpassed in richness, complexity, and the power to absorb.⁴ In what follows, I will survey several aspects of this topic that have received attention so far. My hope is that there will soon be progress both on specific issues and on the general contours of the category. I find many of the puzzles not just intriguing but irksome in their recalcitrance to straightforward analysis.

² One of these connections is to the analysis of discontinuous exceptives such as the French *ne ... que*, on which now see Homer (2015); Authier (2020).

³ This short paper is therefore in the tradition of von Fintel and Kratzer (2014), where we (unsuccessfully) sought to engage the advice of an expert on a set of tricky issues.

⁴ This sentence is a shameless remix of a sentence about “the logic of ordinary speech” from Strawson (1952:p.232).

2.1 *Too and enough*

The founding document of this field is the all too rarely cited Nelson (1980), which I found out about through Humberstone (2006) and Schwarzschild (2008). Nelson discussed structures such as the following:

- (5) This dress is too expensive for me to buy.
- (6) The dress is good enough to wear anywhere.

Nelson identified the comparative nature of the meanings and the presence of hidden possibility modality in the complements of *too* and *enough*. One can easily intuit that (5) means something like “The cost of the dress exceeds any cost at which it would be possible for me to buy the dress”, and that (6) means roughly “The quality of the dress is at least as high as one that would make it possible for it be worn anywhere.” Schwarzschild (2008) proposed a lexical entry for *too* that encodes this kind of meaning, including the hidden modality.

Other work on *too* and *enough* includes Meier (2003); Hacquard (2005); Grano (2022).⁵ Grano tried to adjudicate the question whether the modality is hidden in *too* and *enough* or contributed by the infinitival complement. He provided evidence for the latter. Nevertheless, I suspect there’s more to discover here.

The linguistics of *too* and *enough* is intricate enough to confound even the most sophisticated brains, as shown in the famous puzzler (Wason and Reich (1979); Fortuin (2014)):

⁵ Hacquard (2005) explored the fact that these constructions give rise to the effect of “actuality entailments” that are more well-known in the case of overtly modal constructions. See Grano (2022) for even more references. Meier (2003) added a third expression to the mix: *so ... that* as in *The jet flies so fast that it can beat the speed record*, which is essentially equivalent to *The jet flies fast enough to beat the speed record*. I can’t refrain from mentioning one of my favorite puzzles in linguistics, explored by Hoeksema and Napoli (1993), who discussed the fact that the meaning expressed by *The sun was so hot (that) I fainted* can also be expressed by the paratactic *I fainted, the sun was so hot* (aided by a particular intonational contour). Since the conditional conjunctions to be discussed soon also involve a richer interpretation than one would normally expect in juxtapositions/conjunctions, there may be even more connections to uncover here.

(7) No head injury is too trivial to ignore.

2.2 The SMC

Here's some useful advice for turophiles:

- (8) To get good cheese,
- a. it's enough $\left\{ \begin{array}{l} \text{if you} \\ \text{for you to} \end{array} \right\}$ go to the North End.
 - b. you only have to go to the North End.

Note that we can replace *enough* with the collocation *only have to*. As mentioned in the introduction, this latter construction is the topic of von Fintel and Iatridou (2005a); von Fintel and Iatridou (2007). Follow-up work included Franke (2006); Krasikova and Zhechev (2006); Enguehard (2021); Alonso-Ovalle and Hirsch (2022); Condoravdi and Francez (2022). These works have established that there is something deeply puzzling about how *only have to* gets to express sufficiency. All the accounts attempt something heroic and I'm not convinced by any of them, including our own. Situating the composition of *only have to* in the larger context of enoughology may help in future work on this, so the canvas of possibly related issues and topics that I am painting here can hopefully serve as a map for exploration.

2.3 Connections

2.3.1 Scalar *only*

Several of us have concluded that the *only* in *only have to* has not (just) an exclusive meaning but (also) a merely scalar or mirative impact, on which see, among others, Coppock and Beaver (2014); Alxatib (2020). The core observation here is that there's something odd (or joking) about identifying something remarkable as the "only" thing one has to do to achieve a certain goal:

(9) !!To win the Nobel, you only have to cure cancer.

It appears that this signal is present in lexicalized expressions as well:

(10) !!To win the Nobel, it is $\left\{ \begin{array}{l} \text{sufficient} \\ \text{enough} \end{array} \right\}$ to cure cancer.

2.3.2 Anankastic conditionals

In the course of our investigation, we identified further related constructions that can host the sufficiency meaning. First:

(11) If you want good cheese, $\left\{ \begin{array}{l} \text{it's enough for you to} \\ \text{you only have to} \end{array} \right\}$ go to the North End.

So-called anankastic conditionals such as the one in (11) bring with them a whole other hairball of analytic difficulties, on which see among others: Sæbø (2001); von Fintel and Iatridou (2005b); Huitink (2005a); Huitink (2005b); von Stechow, Krasikova, and Penka (2006); Krasikova (2010); Dunaway and Silk (2014); Condoravdi and Lauer (2016). Most recently, there are Phillips-Brown (2019) and Sæbø (2020), both of whom conclude that anankastic conditionals remain a mystery.

2.3.3 Conditional conjunction

We also pointed out another frame in which the sufficiency modal is at home:

(12) You only have to go to the North End and you'll find plenty of good cheese.

This then would mean engaging with the literature on conditional conjunctions (for a start: Culicover and Jackendoff 1997; Keshet 2012; von Fintel and Iatridou 2017), the core case being something like (13a), which means pretty much the same as the conditional (13b):

- (13) a. I think of him and there are shivers down my spine.
 b. If I think of him, there are shivers down my spine.

Now the SMC-version differs from this core case in two crucial ways: (i) the first conjunct contains the (complex) modal *only have to*, and (ii)

this modal does not appear in any explicit conditional paraphrase: (12) does not mean “if you only have to go to the North End, you’ll find plenty of good cheese”. In fact, the SMC is the only modal that can appear in the first conjunct:

(14)?? You must invest in this company and you will become rich.

Once we’re here, there are more puzzling cases in the vicinity:

(15) a. It won’t take much and she’ll win.

b. It wouldn’t have taken much and she would have won.

2.3.4 Maximizing *all*

Instead of *you only have to go to the North End*, we can also use an *all*-cleft:

(16) To get good cheese, all you have to do is to go to the North End.

On this, see Homer (2019); Tellings (2020).

2.3.5 Sufficiency conditionals

Coppock and Lindahl (2015) discussed another set of cases of minimal sufficiency readings, involving conditionals with some minimizer in the antecedent and variants with a noun phrase in the subject of a causative predicate:⁶

(17) a. If I just think of him, it sends shivers down my spine.

b. Just the thought of him sends sends shivers down my spine.

⁶ Panizza and Sudo (2020) proposed an intricate analysis of the nominal version of this construction without even mentioning the conditional version or discussing the likelihood that the NP stands for something sentential.

2.3.6 Nouwen's puzzle

Finally, we come to a set of observations due to Nouwen (2010b); Nouwen (2010a), namely that the compositional structure of statements of minimal requirements is puzzling:

- (18) The minimum number of points I need to score to win the bet is 300.

Nouwen showed that the meaning of (18) is reached relatively easily if the modal *need* is read as an existential or possibility modal. That of course is not immediately plausible as a meaning for *need*. We come full circle back to Rullmannland: Nouwen suggested *need* here behaves like the Salish modals analyzed in Rullmann, Matthewson, and Davis (2008).⁷ So, one might think that we could make progress here and elsewhere in enoughology if we took into consideration recent work on the quantificational force of modals (Jeretič 2021); Newkirk (2022); Staniszewski (2022).⁸

3 Conclusion

If your head is swirling with all these constructions and the way they might be interconnected, yeah, that's where I am as well. It feels like we have a bunch of ingredients that with some shaking and baking can often give rise to sufficiency meanings. The holy grail is an overarching view that explains what's going on. And I'm afraid I have only scratched the surface. For one thing, all I have talked about is English, but this is of course not (all) parochial to English, as we showed in von Fintel and Iatridou (2005a); von Fintel and Iatridou (2007).⁹ Lastly, I should note that when logicians speak of sufficient (vs. necessary) conditions, it is not clear that we're dealing with the kind of notion of sufficiency (whose foil is excess rather than necessity) we have surveyed here. What's going

⁷ Lassiter (2011) presented an alternative that I am skeptical about.

⁸ Beck (2010) and Dotlačil and Nouwen (2016) contain further relevant discussion in the context of quantifiers in comparatives.

⁹ Fortuin (2013) provided a cross-linguistic survey of ways languages express sufficiency (and excess), without touching on the compositional puzzles I am here concerned with.

on?

So, my plea: Hotze, can you help out an old friend?

References

- Alonso-Ovalle, Luis, and Aron Hirsch. 2022. Keep *only* strong. *Semantics and Pragmatics* 15.
- Alxatib, Sam. 2020. *Focus, evaluativity, and antonymy*. Number 104 in *Studies in Linguistics and Philosophy*. Springer.
- Authier, J.-Marc. 2020. On the comparative analysis of French (*ne*) ... *que* exceptives. *Probus* 32:1–54.
- Beck, Sigrid. 2010. Quantifiers in *than*-clauses. *Semantics and Pragmatics* 3:1–72.
- Beck, Sigrid, and Hotze Rullmann. 1999. A flexible approach to exhaustivity in questions. *Natural Language Semantics* 7:249–298.
- Condoravdi, Cleo, and Itamar Francez. 2022. A minimally sufficient analysis of sufficiency modal constructions. *Amsterdam Colloquium* 23. URL <https://www.dropbox.com/s/qv9qzppggxz23er/Proceedings2022-pages-84-90.pdf>.
- Condoravdi, Cleo, and Sven Lauer. 2016. Anankastic conditionals are just conditionals. *Semantics and Pragmatics* 9:1–69.
- Coppock, Elizabeth, and David I. Beaver. 2014. Principles of the exclusive muddle. *Journal of Semantics* 31:371–432.
- Coppock, Elizabeth, and Anna Lindahl. 2015. Minimal sufficiency readings in conditionals. *Proceedings of the Texas Linguistic Society* 15:24–38. URL <http://www.eecoppock.info/Coppock+Lindahl-TLS15.pdf>.
- Culicover, Peter W., and Ray S. Jackendoff. 1997. Semantic subordination despite syntactic coordination. *Linguistic Inquiry* 28:195–218.
- Dotlačil, Jakub, and Rick Nouwen. 2016. The comparative and degree pluralities. *Natural Language Semantics* 24:45–78.
- Dunaway, Billy, and Alex Silk. 2014. Whither anankastics? *Philosophical Perspectives* 28:75–94.

- Enguehard, Émile. 2021. On the role of alternatives at the semantics-pragmatics interface. PhD thesis, Institut Jean Nicod, École Normale Supérieure. URL <https://theses.hal.science/tel-04021420>.
- von Fintel, Kai, Danny Fox, and Sabine Iatridou. 2014. Definiteness as maximal informativeness. In *The art and craft of semantics*, ed. Luka Crnić and Uli Sauerland, volume 1, 165–174. MIT Working Papers in Linguistics. URL <http://semanticsarchive.net/Archive/jZiNmM4N/FintelFoxIatridou.pdf>.
- von Fintel, Kai, and Sabine Iatridou. 2005a. Anatomy of a modal. In *New work on modality*, ed. Jon Gajewski, Valentine Hacquard, Bernard Nickel, and Seth Yalcin, number 51 in MIT Working Papers in Linguistics. Department of Linguistics and Philosophy, MIT. URL <http://web.mit.edu/fintel/fintel-iatridou-2005-anatomy-wp.pdf>.
- von Fintel, Kai, and Sabine Iatridou. 2005b. What to do if you want to go to Harlem. URL <http://mit.edu/fintel/fintel-iatridou-2005-harlem.pdf>.
- von Fintel, Kai, and Sabine Iatridou. 2007. Anatomy of a modal construction. *Linguistic Inquiry* 38:445–483.
- von Fintel, Kai, and Sabine Iatridou. 2017. A modest proposal for the meaning of imperatives. In *Modality across syntactic categories*, ed. Ana Arregui, María Luisa Rivero, and Andrés Salanova, 288–319. Oxford University Press.
- von Fintel, Kai, and Angelika Kratzer. 2014. Modal comparisons. In *The art and craft of semantics*, ed. Luka Crnić and Uli Sauerland, volume 1, 175–179. MIT Working Papers in Linguistics. URL <http://semanticsarchive.net/Archive/jZiNmM4N/FintelKratzer.pdf>.
- Fortuin, Egbert. 2013. The construction of excess and sufficiency from a crosslinguistic perspective. *Linguistic Typology* 17:31–88.
- Fortuin, Egbert. 2014. Deconstructing a verbal illusion. *Cognitive Linguistics* 25:249–292.
- Franke, Michael. 2006. Teleological necessity and *only*. *Proceedings of the ESSLI Student Session* 11. URL <http://www.sfs.uni-tuebingen.de/~mfranke/Papers/TelNecOnly.pdf>.

- Grano, Thomas. 2022. Enough clauses, (non)finiteness, and modality. *Natural Language Semantics* 30:115–153.
- Hacquard, Valentine. 2005. Aspects of *too* and *enough* constructions. *Semantics and Linguistic Theory (SALT)* 15:80–97.
- Hoeksema, Jacob, and Donna Jo Napoli. 1993. Paratactic and subordinate *so*. *Journal of Linguistics* 29:291–314.
- Homer, Vincent. 2015. *Ne ... que* and its challenges. *West Coast Conference on Formal Linguistics (WCCFL)* 32:111–120. URL <http://www.lingref.com/cpp/wccfl/32/paper3162.pdf>.
- Homer, Vincent. 2019. *That's all*. *West Coast Conference on Formal Linguistics (WCCFL)* 36:1–21. URL <https://www.lingref.com/cpp/wccfl/36/paper3441.pdf>.
- Huitink, Janneke. 2005a. Analyzing anankastic conditionals and sufficiency modals. *Proceedings of ConSOLE* 13:135–156. URL <https://www.hum2.leidenuniv.nl/pdf/lucl/sole/console13/console13-huitink.pdf>.
- Huitink, Janneke. 2005b. Anankastic conditionals and salient goals. *Sinn und Bedeutung* 9:140–154.
- Humberstone, Lloyd. 2006. Sufficiency and excess. *Proceedings of the Aristotelian Society: Supplemental Volumes* 80:265–320.
- Jeretič, Paloma. 2021. Neg-raising modals and scaleless implicatures. PhD thesis, New York University.
- Keshet, Ezra. 2012. Focus on conditional conjunction. *Journal of Semantics* 30:211–256.
- Krasikova, Sveta. 2010. Sufficiency inference in anankastic conditionals. *Semantics and Linguistic Theory (SALT)* 20:91–108.
- Krasikova, Sveta, and Ventsislav Zhechev. 2006. You only need a scalar *only*. *Sinn und Bedeutung* 10:199–209.
- Lassiter, Daniel. 2011. Nouwen's puzzle and a scalar semantics for obligations, needs, and desires. *Semantics and Linguistic Theory (SALT)* 21:694–711.
- Meier, Cécile. 2003. The meaning of *too*, *enough*, and *so...that*. *Natural*

Language Semantics 11:69–107.

- Nelson, Eric S. 1980. *Too and enough*. *Minnesota Papers in Linguistics and the Philosophy of Language* 6:93–132.
- Newkirk, Lydia. 2022. Be flexible, but not too flexible: Limited variable-force modals in Kinande and the typology of modal force. PhD thesis, Rutgers University.
- Nouwen, Rick. 2010a. Two kinds of modified numerals. *Semantics and Pragmatics* 3:1–41.
- Nouwen, Rick. 2010b. Two puzzles about requirements. In *Logic, language and meaning*, ed. Maria Aloni, Harald Bastiaanse, Tiki de Jager, and Katrin Schulz, number 6042 in *Lecture Notes in Computer Science*, 345–354. Springer.
- Panizza, Daniele, and Yasutada Sudo. 2020. Minimal sufficiency with covert *even*. *Glossa: a journal of general linguistics* 5.
- Phillips-Brown, Milo. 2019. Anankastic conditionals are still a mystery. *Semantics and Pragmatics* 12:1–17.
- Rullmann, Hotze, Lisa Matthewson, and Henry Davis. 2008. Modals as distributive indefinites. *Natural Language Semantics* 16:317–357.
- Schwarzschild, Roger. 2008. The semantics of comparatives and other degree constructions. *Language and Linguistics Compass* 2:308–331.
- Staniszewski, Frank. 2022. Modality and time in logical context. PhD thesis, Massachusetts Institute of Technology.
- von Stechow, Arnim, Sveta Krasikova, and Doris Penka. 2006. Anankastic conditionals again. In *A Festschrift for Kjell Johan Sæbø*, ed. Torgrim Solstad, Atle Grønn, and Dag Haug, 151–171.
- Strawson, P.F. 1952. *Introduction to logical theory*. Methuen.
- Sæbø, Kjell Johan. 2001. Necessary conditions in a natural language. In *Audiatur vox sapientiae*, ed. Caroline Féry and Wolfgang Sternefeld, 427–449. Akademie Verlag.
- Sæbø, Kjell Johan. 2020. Anankastic conditionals. In *The Wiley Blackwell companion to semantics*, ed. Daniel Gutzmann, Lisa Matthewson, Cécile Meier, Hotze Rullmann, and Thomas Ede Zimmermann.

Tellings, Jos. 2020. An analysis of *all*-clefts. *Glossa: a journal of general linguistics* 5.

Wason, Peter C., and Shuli S. Reich. 1979. A verbal illusion. *Quarterly Journal of Experimental Psychology* 31:591–597.

A sketch of content question formation in Eauripik Woleaian*

LAURA GRIFFIN
University of Toronto

1 Introduction

While many languages have content-based questions, they differ widely in terms of potential movement and extraction out of the relevant phrase. Research on content questions in Austronesian languages largely focuses on the usage of pseudoclefts in both predicate-initial languages, such as Malagasy (Potsdam 2006, 2009) and Fijian (Potsdam 2009), and subject-initial languages, such as Marshallese (Willson-Sturman 2014). I contribute novel data on Woleaian, a Chuukic language spoken in the Federated States of Micronesia, which appears to have a movement and non-movement strategy for content question formation. Crucially, I propose that one of these movement strategies involves clefting rather than pseudoclefting. In this paper, I present a description of content questions in the Eauripik dialect of the Woleaian language, challenging previous literature stating that there are only *in situ* content questions in the language.

2 Background literature

The Woleaian language (ISO 639-3: woe), also known as *kapetele faliuwash*, is spoken by approximately 2,000 native speakers in Yap State in the Federated States of Micronesia. Woleaian is an Austronesian language of the Chuukic subgroup that has been influenced by Japanese, English, and Spanish (Sohn 1975:1). Woleaian is related to neighboring languages such as Ulithian, Satawalese, and Chuukese (Sohn 1975:4).

* Author contact information: laura.griffin@mail.utoronto.ca. I would like to extend my sincere thanks to Lenny Saumar for sharing his language and his time so generously, and without whom this research would be absolutely impossible. I also thank Audrey Lai and Clemens Mayer for all of their help and patience, Shinichiro Fukuda and Bradley McDonnell for their guidance, and the Austronesian Circle for all of their helpful comments and feedback.

I would finally like to thank Hotze Rullmann, my very first semantics teacher and the inspiration behind my forays into tense, aspect, and the syntax-semantics interface. Thank you for your endless patience, good humour, and thought-provoking questions!

The primary documentation of this language was done by Ho-min Sohn and Anthony F. Tawerilmang. Data collection occurred from 1971–1975 (Sohn & Tawerilmang 2019:vii). These two scholars produced both a reference grammar (1975), an English-Woleaian dictionary (2019), and an orthography based on the dialect spoken in the Woleai atoll. The orthography used in the grammar and dictionary is currently outdated, as the orthography was modified after publication in the 1990s, but is still understandable to speakers today (Lenny Saumar, p.c.).

There are approximately six language varieties identified by Sohn (1975:5): Woleai, Eauripik, Faraulep, Elato, Lamotrek, and Ifaluk. The Woleai variety is further divided into the East and West dialect. The dialect under focus in this paper is the variety spoken in the Eauripik atoll by one speaker in his early 30s. While there is no comprehensive study of how these dialects differ from one another, there are syntactic or phonological differences between the data presented here and that in Sohn (1975), which could be licensed by language change, dialectal variation, or a combination of both.

Woleaian is a canonically SVO language for both transitive (1) and intransitive (2) contexts.

- (1) Ye mongo iige.
 3SG.A eat fish
 ‘S/he eats fish.’¹
- (2) Go tefale.
 2SG.A return
 ‘You return.’

A summary of attested agreement markers is reproduced in Table 1 from Mayer (in prep.). The agentive forms have both bound and unbound variants, though this variation seems likely due to orthography rather than the language grammar. Argument morphology appears to be predicate-dependent, as some predicates require both overt subject and object agreement, while others only require one or none.

¹ Data are presented using the current Woleaian orthography. Glosses used in this paper: A = agentive, ANA = anaphor, DEM = demonstrative, FOC = focus, MED = medial, P = patient, PFV = perfective, PL = plural, POSS = possessive, SG = singular, and VIS = visible. All language data were provided by Lenny Saumar unless otherwise noted.

Table 1: Argument marking in Woleaian, Eauripik dialect (Mayer, in prep.)

	Agentive	Patient	Emphatic
1SG	<i>ie/i=</i>	<i>yei</i>	<i>gaang</i>
2SG	<i>go/go=</i>	<i>=go</i>	<i>geelee</i>
3SG	<i>ye/ye=</i>	<i>=i, =we</i>	<i>ie</i>
1PL.INCL	<i>si/si=</i>	<i>geshe</i>	<i>gishe</i>
1PL.EXCL	<i>gai</i>	<i>gamem</i>	<i>gamem</i>
2PL	<i>gai</i>	<i>gami</i>	<i>gami</i>
3PL	<i>re/re=</i>	<i>re</i>	<i>ire</i>

Lastly, Woleaian has overt focus marking that can co-occur with other elements, such as demonstratives in (3). Capital letters in the English translation are used to mark the focussed element. This focus marker marks number by adding the plural morpheme *ka*, as in (4). These markers may also function as anaphors and are not limited to occurring only in focus constructions.

- (3) Gelaago we mene ye gangi.
 dog DEM FOC.SG 3SG.A eat
 ‘The DOG ate the food.’

- (4) Gelaago kawē meka re gangi.
 dog DEM.PL FOC.PL 3PL.A eat
 ‘The DOGS ate the food.’

Sohn (1975) describes a similar word *mele* as a ‘selective emphasis marker’ (1975:175), as in (5). This word is derived from the base *mel-*, an anaphoric referent marker. This marker can take *-le*, the third person singular possessive marker, to derive *mene*, the focus marker.²

- (5) Metta mele go weri?
 What FOC 2SG.A see
 ‘What did you see?’ (Sohn 1975:176)

² I have added interlinear glosses to all examples from Sohn (1975) in accordance with definitions from Sohn and Tawerilmang (2019) and my own analysis of the data; all errors in glossing are mine.

Sohn (1975) has a short description dedicated to questions in the language. Polar questions are described as structurally equivalent to their declarative counterparts, as (6), a question, and (7), a statement, have completely identical lexemes and morphemes. They differ only in intonation. Whereas declarative sentences have a flat mid-high to low pitch, polar questions have a very high to mid-low pitch sentence-finally (Sohn 1975:39, 41). Content questions have a similar pitch contour to declarative sentences (Sohn 1975:40).

(6) Ye sa lag?
 3SG.A PFV go
 ‘Did he go?’ (Sohn 1975:154)

(7) Ye sa lag.
 3SG.A PFV go
 ‘He went.’ (Sohn 1975:91)

Content questions, also known as *wh*-questions, are formed through the use of a content question word in an interrogative context. They are described as only occurring *in situ*, along with other information about their restrictions in equational constructions. This conclusion does not appear to be supported in Eauripik Woleaian, where content question phrases are not limited to only occurring in their base-generated position.

Examples of the content question word’s *in situ* status are presented in the following sentences including the words *iteiu* ‘who’, *ileet* ‘when’, and *iiya* ‘where’ (8a–c). Sohn (1975) includes examples of content questions with the focus marker *mele* but does not explain its occurrence or distribution.

(8) a. Iteiu mele ye weri?
 who FOC 3SG.A see
 ‘Who is the one who saw it?’ (Sohn 1975:169)

b. Re sa lag ileet?
 3PL.A PFV go when
 ‘When did they go?’ (Sohn 1975:169)

c. John ye bel lag iiya?
 John 3SG.A will go where
 ‘Where will John go?’ (Sohn 1975:169)

In Eauripik Woleaian, polar questions are also structurally equivalent to content questions and differ primarily in prosody. Polar and content questions may optionally occur with the question particle *go*, which causes a rising intonation. Examples (9a) and (9b) show no difference in morphology, apart from the optional question particle. Both (9b) and (9c) may occur with *go* without any change in grammaticality.

- (9) a. Go gabiungiu kapete-le faliu-wash.
 2SG.A teach language-3SG.POSS island-1PL.POSS
 ‘You teach Woleaian.’
 (Lit. ‘You teach the language of our island.’)
- b. Go gabiungiu kapete-le faliu-wash (go)?
 2SG.A teach language-3SG.POSS island-1PL.POSS (QP)
 ‘Do you teach Woleaian?’
 (Lit. ‘Do you teach the language of our island?’)
- c. Iteoiu ye gabiungiu kapete-le
 who 3SG.A teach language-3SG.POSS
 faliu-wash (go)?
 island-1PL.POSS (QP)
 ‘Who teaches Woleaian?’

3 Methodology

These data were collected through elicitations with a consultant, Lenny Saumar, in a field methods class from September 2021 to May 2022 at the University of Hawai‘i at Mānoa and in follow-up online elicitations afterwards in late 2022.³ Lenny is a 34-year-old native speaker of the Eauripik variety of Woleaian and an L2 speaker of English and Chuukese. He spent his childhood on the Eauripik atoll, where approximately 100 people reside. He currently resides in O‘ahu, Hawai‘i, where he uses the language daily with friends and family.

Two main tasks were conducted during elicitations. The consultant was either presented with an English sentence and context and asked to translate from English to Woleaian or given an English context and Woleaian sentence and asked to judge the felicity of the sentence in the

³ The IRB proposal for this project is 2021-00641 through the University of Hawai‘i at Mānoa and has an expiry date of September 12, 2021. Any questions about the protocol may be directed to Dr. Bradley McDonnell, the principal investigator of the project.

provided context. Additional tasks included verifying the grammaticality of sentences from Sohn (1975), recounting a traditional narrative, and identifying images of plants and animals to create a variety of naturally occurring to more controlled speech.

4 Data

Eight primary content question words were identified throughout the course of elicitations. Table 2 presents a brief overview of each word in Woleaian, presented in the orthography, and an approximate English translation. I will next focus on the distribution of two content question words, *meta* ‘what, how’ and *iteoiu* ‘who’ as examples of question formation strategies in Eauripik Woleaian.

Table 2: Content question words in Woleaian, Eauripik dialect

WH Word	English Gloss	WH Word	English Gloss
<i>Meta</i>	‘what, how’	<i>Fitou</i>	‘how many’
<i>Iteoiu</i>	‘who’	<i>Ifa/Ikefa</i>	‘which’
<i>Ileete</i>	‘when’	<i>Fasengale</i>	‘how’
<i>Iya/iiya</i>	‘where’	<i>Meta faale</i>	‘why’

Based on the data collected, there appear to be three strategies used to form questions: *in situ*, cleft movement, and focus movement. These strategies are syntactically and pragmatically restricted, depending on the type of extraction and on the surprise of the speaker.

4.1 In situ questions

In situ questions are the primary content-question formation strategy described by Sohn for the Woleai dialect of Woleaian. These questions involve no overt movement of the relevant phrase from where it was base-generated, as in (10) and (11). This strategy can be used for subject, object, and adjunct positions for *iteoiu* ‘who’, as in examples (10a–c), and for object and adjunct positions for *meta* ‘what, how’, as in (11a–b), but this strategy is ungrammatical for *meta* in subject position, as in (11c).

- (10) a. Iteoiu gangi mongo?
 who eat food
 ‘Who eats food?’

- b. Go shungali iteoiu?
 2SG.A meet who
 ‘Who do you meet?’
- c. Go mongo fituge re-li iteoiu?
 2SG.A eat meat with-3SG who
 ‘Who do you eat meat with?’
- (11) a. Ye gangi meta?
 3SG.A eat what
 ‘What did s/he eat?’⁴
- b. Go gangi ngali meta?
 2SG.A eat with what
 ‘What are you eating with?’
- c. *Meta pwule?
 what burn
 Intended: ‘What burns?’

4.2 Cleft constructions

The first of the two movement options involves creating a cleft construction as part of the content question. To diagnose these constructions, I follow diagnostics put forth in Potsdam and Polinsky (2011). Displacement or true movement structures have content question phrases that lack predicative properties, are monoclausal, and have an “activated left periphery” that allows for multiple constituents (Potsdam & Polinsky 2011:119, 121). In contrast, pseudoclefts and clefts have a content question phrase as their main predicate and a relative clause as the remainder of the construction, two hallmarks of this biclausal construction (Potsdam & Polinsky 2011:113). Pseudoclefts and clefts differ in their treatment of the construction’s remainder, where in pseudoclefts they exhibit nominal and subject properties similar to those in relative clauses. In addition, pseudoclefts may have a ‘dummy’ or expletive head present in the remainder clause, whereas clefts do not.

⁴ Woleaian does not have overt tense marking (Sohn 1975:233) and instead marks temporal reference through aspect. The tense provided in translations is largely based on context.

Clefts allow pied-piping in the content question phrase, but pseudoclefts do not.

Examples of a cleft question are given for both *iteoiu* ‘who’ in (12a) and *meta* ‘what, how’ in (12b). These constructions involve the content question word occurring on the left periphery, regardless of where it was base generated, followed by a demonstrative marker and then the rest of the predicate.

- (12) a. Iteoiu i-ka ie shungi-re?
 who ANA-PL 1SG.A meet.PL
 ‘Who (pl.) is it that I met?’
- b. Meta uru ye go rongi?
 what song DEM 2SG.A sing
 ‘What is the song that you sang?’

Firstly, dummy heads are possible in headless relative clauses. The dummy head *gena* ‘person’ can serve as the head for (13), replacing the demonstrative *ye*. This suggests that demonstratives can behave as a dummy head that lacks semantic value but satisfies markedness constraints.

- (13) Re-li iteoiu gena go mongo fituge la?
 with-3SG who person 2SG eat meat MED.VIS.DEM
 ‘With which person did you eat meat?’

Secondly, content questions that are extracted out of adjunct position allow for pied-piping, as in (14a). The content question word and its preposition may also be separated, as in (14b). Potsdam and Polinsky (2011) argue that PPs are generally not pied-piped in Austronesian languages, which is evidence for a pseudocleft construction. However, in (13), both pied-piping and a dummy head can co-occur. Based on the presence of the dummy head, I propose that Eauripik Woleaian has content question pseudoclefts rather than clefts.

- (14) a. Re-li iteoiu ye go mongo fituge?
 with-3SG who DEM 2SG.A eat meat
 ‘With whom did you eat meat?’
- b. Iteoiu ye go mongo fituge re-li?
 who DEM 2SG.A eat meat with-3SG

‘Who did you eat meat with?’

Pseudocleft constructions are attested for subject, object, and adjunct extraction for *iteoiu* ‘who’, as in (15a–c) and subject, object, and adjunct extraction for *meta* ‘what, how’, as in (16a–c).

- (15) a. Iteoiu ye gangi mongo?
 who 3SG.A eat food
 ‘Who is it that eats food?’
- b. Iteoiu ye go shungali?
 who 3SG.A 2SG.A meet
 ‘Who is it that you meet?’
- c. Iteoiu iye go kauru-re movie we reli?
 who DEM 2SG.A watch-3PL.P movie DEM with
 ‘Who did you see the movie with?’
- (16) a. Meta ye pwule?
 what 3SG.A burns
 ‘What is it that burns?’
- b. Meta ye gangi?
 what 3SG.A eat
 ‘What is it that s/he eats?’
- c. Meta iye go gangi ngali?
 what DEM 2SG.A eat with
 ‘What is it that you eat with?’

4.3 Focus movement

The last movement strategy available for content question formation involves focus movement. There are two focus markers, *mene* for singular constituents and *meka* for plural constituents. These constructions involve the content question word or phrase occurring on the left periphery, again regardless of where it was base generated, followed by the focus marker and then the rest of the predicate. Content questions with focus movement are attested for subject, object, and adjunct extraction for *iteoiu* ‘who’ and subject, object, and adjunct extraction for *meta* ‘what, how’. For the subject extraction cases in (17a)

and (18a), focus movement triggers a subject pronoun to occur in the predicate, following the focus marker and preceding the verb. I leave this for future investigation to determine whether focus-marking in subject extraction cases triggers a dummy subject or whether focus-marking co-occurs with cleft constructions.

- (17) a. Iteoiu mene ye gangi mongo?
who FOC.SG 3SG.A eat food
'Who is eating food?'
- b. Iteoiu mene go tuguwe?
who FOC.SG 2SG.A punch
'Who did you punch?'
- c. Iteoiu mene go mongo fituge re-li?
who FOC.SG 2SG.A eat meat with-3SG
'Who do you eat meat with?'
- (18) a. Meta mene ye pwule?
what FOC.SG 3SG.A burn
'What is it that burns?'
- b. Meta mene ie gangi?
what FOC.SG 1SG.A eat
'What did I eat?'
- c. Meta mene ie mongo fituge ngali?
what FOC.SG 1SG.A eat meat with
'What did I eat the meat with?'

The focus markers *mene* and *meka* appear to mark contrastive focus rather than informational focus: instead of introducing new information, the markers highlight that one option rather than another has been selected. Focus markers are subject to markedness restrictions: they cannot occur sentence-finally, as in (19b) and (20b), for both interrogative and declarative contexts.

- (19) a. Meta mene ye gangi?
what FOC.SG 3SG.A eat
'What did s/he eat?'

- b. * Ye gangi meta mene?
 3SG.A eat what FOC.SG
- (20) a. Iige mene ye mingo.
 fish FOC.SG 3SG.A eat
 ‘S/he ate fish.’
- b. *Ye mingo iige mene.
 3SG.A eat fish FOC.SG

The focus markers also cannot remain *in situ* even if they do not occur sentence-finally, as in (21b). Due to these restrictions, it appears that these focus markers license movement to the left periphery of the phrase.

- (21) a. Meta meka go chuwai me stowa?
 what FOC.PL 2SG.A buy from store
 ‘What (pl.) did you buy at the store?’
- b. *Go chuwai meta meka me stowa?
 2SG.A buy what FOC.PL from store

5 Pragmatic implications

The three content question-forming strategies (*in situ*, pseudoclefted, and focused) not only differ in their syntax, but also in their pragmatics.

- (22) a. Meta ye gangi?
 what 3SG.A eat
 ‘What is it that s/he ate?’
- b. Meta mene ye gangi?
 what FOC.SG 3SG.A eat
 ‘What did s/he eat?’
- c. Ye gangi meta?
 3SG.A eat what
 ‘What did s/he eat?’

This set of data is organized from most to least pragmatically neutral. The first sentence with the pseudocleft construction, (22a), is the most neutral reading, where the speaker has full knowledge of all foods

available. The second sentence with focus movement, (22b), has a slight reading of surprise, where the speaker did not know all foods that were available. The last sentence, (22c), has a strong surprise interpretation, where everything is new information. This sentence can also have an echo question reading, such as asking for clarification, but does not have an intensification interpretation, such as a ‘What the hell?’ reading.

6 Conclusion

Euaripik Woleaian appears to have two main strategies to create content questions, one that involves the content question word remaining *in situ* and one that involves movement. This differs from previous literature, which did not include movement as a possible strategy. I propose that the movement strategy is decomposable to two main constructions involving focus movement and pseudoclefting constructions. These data supplement existing documentation, providing a new analysis of interrogative formation that can be used as a stepping point for future work. This description lends itself to supporting future investigations into island effects, resumption, and relative clauses, as well as the relationship between focus and demonstrative marking in both interrogative and declarative sentences. Overall, this research adds to the growing body of literature on question-formation strategies in subject-initial Austronesian languages.

References

- Mayer, Clemens J. (in prep.). *A Grammar of Woleaian*. Ph.D. dissertation, Department of Linguistics at the University of Hawai‘i at Mānoa.
- Potsdam, Eric. 2006. The Cleft Structure of Malagasy Wh-Questions. In H. Gärtner, P. Law & J. Sabel (eds.), *Clause Structure and Adjuncts in Austronesian Languages*. Berlin/Boston: De Gruyter Mouton, 195–232.
- Potsdam, Eric. 2009. Austronesian verb-initial languages and wh-question strategies. *Natural Language & Linguistic Theory* 27(4):737–771.
- Potsdam, Eric & Maria Polinsky. 2011. Questions and Word Order in Polynesian. In Claire Moyse-Faurie and Joachim Sabel (eds.), *Topics in Oceanic Morphosyntax*. Berlin/Boston: De Gruyter Mouton, 107–134.

- Sohn, Ho-Min. 1975. *Woleaian Reference Grammar*. Honolulu: University of Hawaii Press.
- Sohn, Ho-Min & Anthony F. Tawerilmang. 1976. *Woleaian-English Dictionary*. Honolulu: University of Hawaii Press.
- Willson-Sturman, Heather. 2014. Pseudocleft constructions in Marshallese wh- questions. *Lingua* 145:1–35.

If I had half a brain: Polarity sensitive idioms in conditional clauses

JACK HOEKSEMA
University of Groningen

1 Introduction

Polarity items have distributional limitations which have puzzled linguists ever since Klima (1964) came up with the first formal and explicit account of English *any* and its ilk. A class of environments that has been recognized from the start as relevant to the analysis of polarity items is that of conditional clauses, more precisely the protasis of a conditional. Polarity items may appear there without the support of negation (in spite of the fact that they do fail the tests for downward entailment; cf., e.g., von Stechow 1999). In the following examples, a sequence of conditional clauses provides a home for a variety of polarity items, rendered in boldface for easy detection:

- (1) If it helps **any**, I love you. [Episode of *Love Boat*]
- (2) If I get my car back **any** different than I gave it, Monster Joe's gonna be disposing of two bodies.
[from *Pulp Fiction*, Quentin Tarantino, 1994]
- (3) If **ever** a man could wheedle his way into a wench's affections, it was Edgar. [Michael Jecks, *The Bishop Must Die*]
- (4) You don't comply with the conditions if you **budge** from the office during that time.
[Sir Arthur Conan Doyle, *The Adventures of Sherlock Holmes*]

The above examples, all involving clauses introduced by *if*, are pragmatically diverse. Example (1), for instance, is a biscuit conditional (cf. Austin 1956; Rawlins 2020), example (2) would count as a threat (cf. Lakoff 1969; Czipak 2014 for discussion of the difference between threats and promises with regard to polarity licensing), and examples (3) and (4) are neither. Counterfactual conditionals may also host polarity items:

- (5) Had he **ever** been in the way of learning, I think he would have drawn very well. [Jane Austen, *Sense and Sensibility*]
- (6) If **anyone** had **ever** replaced the top layer of blotting paper, Solly's gym would have ground to a halt. [Bryce Courtenay, *The Power of One*]

The protasis part need not have the form of a regular conditional clause:

- (7) You **lay a hand on** her **ever** again, and I will take you out. [Episode of *Nash Bridges*]
- (8) You **breathe a word** of this to Buffy and I'll see to it that you end up in the ground. [Episode of *Buffy the Vampire Slayer*]
- (9) What do you expect to gain from seeing Sarah-Jane Beckett, assuming you can **even** find her? [Elizabeth George, *A Traitor to Memory*]

Presumably cases such as (10), with an infinitival subject as host, can be viewed as belonging to the conditional supercategory as well, given their conditional interpretation, as shown by the possibility of paraphrases such as (11), and their dependence on modal elements such as *will* or *may* (compare 12).

- (10) Measuring a particle, or disturbing it **in any way**, will cause the superposition to “decohere” or collapse. [*Washington Post*, October 23, 2019]
- (11) The superposition will “decohere” if you measure or disturb a particle in any way.
- (12) Measuring a particle (*in any way) caused the superposition to “decohere” or collapse.

An interesting and rather problematic case is that of conditional *with/without* (discussed in Reuneker 2016; Hoeksema 2022):

- (13) With **any** luck, we will be on time.
[= If we have any luck, we will be on time.]

(14) Without **some** luck, we won't be on time.

In a corpus search for polarity items licensed by conditional *with*, only the string *with any luck* stood out. Other cases that might seem feasible do not ring true:¹

(15) ? With any booze, we can have a party.

(16) ? With any money, we should buy a car.

(17) ? With any books, we could start reading.

So this looks like a lexically restricted pattern. With non-polarity items, conditional readings are more easily available, both in English and in Dutch. Compare for example:

(18) With some more booze, this party could get a lot wilder.

(19) With your talent, I would have made captain at 25.

2 Polarity items that shun conditionals

Not all polarity items were created equal. Whereas well-known items such as *any* and *ever* are generally accepted in conditional clauses, various others are not. The Dutch modal auxiliary *hoeven* and its German and English counterparts *brauchen* and *need* are not licit in conditional clauses (Zwarts 1981; van der Wouden 2001). The following examples may illustrate this (they are not glossed but all have the same meaning):

- (20) a. * Please let us know if you need eat.
b. * Laat het ons a.u.b. weten als u hoeft eten.
c. * Bitte lassen Sie uns wissen, wenn Sie zu essen brauchen.

Note that the semi-auxiliary *need* (which is followed by *to*) is not polarity-sensitive, and may appear in all sorts of contexts, including conditional clauses, without any licensing elements:

¹ Marcin Morzycki (p.c.) pointed out that the examples in (15) to (17) sound better when *at all* is added. This seems like a case of parasitic licensing (Den Dikken 2006; Hoeksema 2007): In some cases, licensing of a polarity item is made possible by the addition of another polarity item.

- (21) a. We need to eat.
b. Please let us know if you need to eat.

An interesting exception to the ungrammaticality of cases such as (20a) is provided by the fixed expression *if need be*.

Hoeksema (2012) also lists *anymore* as shunning conditionals (in varieties of English which lack positive *anymore*), and the class of temporal expressions exemplified by *in ages*, *in years*, *in decades*, *in weeks* etc. Compare:

- (22) *Let me know if you have seen her in ages.

The gap does not seem to be random. Not only does it hold for the entire set of *in X* expressions, it also applies to their Dutch counterparts:

- (23) *Als je haar in tijden hebt gezien, laat het me weten.
if you her in times have seen let it me know
'If you have seen her in ages, let me know.'

3 Averidical conditionals

Rullmann (2003:349) mentions *either* as a polarity item that fails to appear in conditional clauses. However, he cites an example by Larry Horn (p.c.) that is fully acceptable:

- (24) I didn't take semantics. I'll be damned if I take pragmatics, **either**.

Horn (1989:348) gives similar examples with *until*. Let's call conditionals such as the above *averidical*. A note on terminology: Averidical conditionals are to be distinguished from counterfactual conditionals. They are not about what might have happened under different circumstances. In particular, the apodosis is not to be taken literally. Taking or not taking a class in pragmatics is no cause for eternal damnation, as far as I can tell, based on the limited information on this subject in the Bible. Rather, *be damned if* has an additional, idiomatic use in which it expresses a negative intention. In its most typical use, it is first person: *I'll be damned if I do that = I won't do that*. Third person use is possible too, provided the perspective of the third person is taken, for instance in reported speech (*He said he would be damned if he took another semantics class*) or free indirect style. Since this use appears to be connected to the formulation of intentions, it is future-oriented. It may

be in the past tense, but only to express past intentions: *He would be damned if he took another class with Professor Rullmann, the student muttered.* Note that we should not confuse averidical *be damned if* with another frequent idiomatic use: *You will be damned if you do and damned if you don't*, which is reserved for situations in which there is no attractive path forward. Intentions play no role here. The latter use is fine with second person pronouns (generic or otherwise), unlike the former use. After all, it is pretty strange to tell your interlocutor what their intentions are.

Besides negative intentions (by far the most common case), epistemic interpretations may also be involved sometimes with averidical conditionals. The Cambridge online dictionary states for the expression *be hanged if* the following: “used to express your determination not to do something or not to allow someone else to do something.” In addition to this, however, it also lists the idiom *I'll be hanged if I know*, which it describes as being “used to say that you certainly do not know.” The epistemic state of not knowing is presented as a certainty, by using the common assumption that people do not fancy being hanged, and so would only offer that option if they knew it to be vacuous.

Averidical conditionals are idiomatic and cannot be freely formed. E.g., *I will be dead if I know what you mean* or *I will swallow poison if I take semantics* do not strike me as acceptable alternatives to the *damned/hanged* cases. It is also worth pointing out that the order of protasis and apodosis is fixed. The following example, while grammatical, appears to have a literal reading only:

(25) If I take another semantics class, I'll be hanged.

Dutch has a fairly wide variety of idioms that serve in averidical conditionals. Here is a list of cases I have encountered:

- | | | | |
|------|----|-------------------|------------------|
| (26) | a. | mogen doodvallen | ‘may drop dread’ |
| | b. | mogen hangen | ‘may hang’ |
| | c. | een boon zijn | ‘be a bean’ |
| | d. | mijn kop eraf | ‘my head off’ |
| | e. | mogen barsten | ‘may burst’ |
| | f. | mogen sterven | ‘may die’ |
| | g. | mogen doodsmakken | ‘may drop dead’ |
| | h. | zullen liegen | ‘would lie’ |
| | i. | zijn muts opeten | ‘eat one’s cap’ |

Some of these are similar to the English *be damned if/be hanged if*, involving various unsavory ways of dying, whereas others present a ludicrous, impossible state of affairs, such as being a bean or eating one's hat. In this group we may also place expressions of the form *mijn naam is geen X* 'my name is not X', where X is the actual name of the speaker. The following example is of special interest due to the presence of the polarity item *pluis* 'okay, safe':

- (27) Als dat **pluis** is, dacht de man, dan heet ik
if that okay is thought the man then be.called I
geen Japik meer.
no Japik anymore
'If that is in order, the man thought, my name is no longer Japik.'

The appearance of *pluis* is interesting, because this expression has a rather limited distribution. Van der Wouden (1994) treats it as a superstrong NPI, licensed only by regular negation, not even by n-words, but the above example shows it to be licit in averidical conditionals as well.

4 Special idiomatic cases of NPIs in conditionals

Part of Rupert Holmes' Piña Colada Song (1979) goes like this:

- (28) If you're not into yoga
If you have half a brain
[..]
Write to me, and escape.

Such examples beg for an analysis in terms of a hidden 'even' (popular in the NPI literature, cf. Heim 1984, Lee & Horn 1994, Rullmann 1996, Guerzoni 2003, Crnić 2019, among others, for discussion and various proposals). The speaker is not looking for someone with only half a brain, but for someone smart enough to escape. And someone who is not into yoga and has at least half a brain, should realize this. (Apologies to all smart people who are into yoga.) As an NPI, *half a brain* is interesting because it appears only in relative clauses modifying universal quantifiers and negative quantifiers (*anyone with half a brain*, *nobody with half a brain*) and in conditional clauses. Regular negation is out of the question:

- (29) *Fred did not have half a brain.

Most accounts of polarity items, including those of Zwarts (1998), Giannakidou (2011), and Gajewski (2011), have problems with such a distribution, since they all assume that any polarity item, be it weak, strong, or whatever, may be licensed by negation. If we assume, however, that some polarity items may have additional requirements, apart from polarity licensing, perhaps we could treat *half a brain* as such, and would not have to burden the theory of polarity licensing even further. The expression seems to function as a restriction of a set of persons to those individuals that have at least a minimal level of intelligence. In sentences such as (29), *half a brain* does not seem to have this purpose. That we are dealing with a minimal requirement, not a maximal requirement, may be illustrated by pairs such as the following:

- (30) a. This should be obvious to anyone with half a brain.
b. # This will be hard to grasp for anyone with half a brain.
- (31) a. If you have half a brain, you will grab this opportunity.
b. # If you have half a brain, you won't be able to grasp this opportunity.
- (32) a. Nobody with half a brain will feed the polar bears.
b. # Nobody with half a brain will figure out how to feed the cat.

The b-sentences are not ungrammatical but have a literal reading. This reading will not be impacted if we add *at most to half a brain*. The a-sentences, on the other hand, are best understood as implying *at least half a brain*.

A type of expression that appears to be completely restricted to conditional clauses is exemplified by the following sentences:

- (33) If the past month is any guide, it is the more freewheeling films that are likelier to be box-office hits.
[*The Economist*, July 21, 2018, p. 44]
- (34) If history is any guide, the coronavirus's impact on the poor will be felt long after the pandemic is over.
[*The New York Times*, August 3, 2020]
- (35) If his general performance is anything to go by, I'd say he got it wrong.
[*The Guardian*, December 7, 2010]

- (36) If the unpredictable London weather is anything to go by, you might be needing a dose of that right about now.

[*Evening Standard*, May 19, 2023]

For *any* in general, about 5% of all occurrences are in conditional clauses, slightly more if you discount free choice cases. (There is considerable variation between text types, so these figures do not mean much, except that they are fairly low.) Now it is curious to see that *any guide* when used as a predicate nominal, as in (33) and (34), has 100% occurrences in conditional clauses. Much the same is true of *anything to go by*, a virtual synonym of *any guide*. Other contexts feel odd:²

- (37) #I do not believe that history is any guide.

- (38) #Nothing from the past is any guide.

- (39) #No politician is any guide.

- (40) #Very few things are any guide.

- (41) #Was history ever any guide?

- (42) #Little else is any guide.

If you replace *guide* by *good*, all of the above will be just fine. While *be any guide* or *be anything to go by* do not seem to be regular idioms (their interpretation is compositional), they seem to exemplify a general schematic pattern *if X is any guide, then Y*, where Y is some proposition whose validity is based on generalizing from X. The truth of Y is not dependent on that of the protasis, but rather, the protasis seems to be a hedge, warning about the limited validity of Y.

5 Conclusions

In this brief paper we have encountered expressions which should not, but do as a matter of fact, show up in conditional clauses (*with any luck*, *if need be*), items which should, but do not, appear in conditional clauses

² Marcin Morzycki does not fully share my aversion to (37) to (42). Judgments may be tenuous.

(*in years, in ages*), items which appear in averidical conditionals and no other conditionals (*either*).

We took a detour to consider the idioms that make up the core of averidical conditionals and noted some interesting subtypes that could be studied in more detail. One subtype is connected to negative intentions; another has to do with absence of knowledge. In Dutch, the cases involving undesirable ways of dying (*Ik mag doodvallen als ik dat ga doen* ‘I may drop dead if I am going to do that’) mostly involve the former type, whereas predicates ascribing impossible properties to individuals appear to favor statements expressing lack of knowledge or understanding (*Ik ben een boon als ik dat snap* ‘I am a bean if I understand that’). Are there more subtypes, and how are they connected to various averidical conditionals in English? Could this be a topic for crosslinguistic research?

Point of view issues were briefly mentioned. Averidical conditionals as well as ascriptions of knowledge require a point of view. Often, they are first person. If not, they can be third person, requiring us to take the perspective of that person. Second person cases are pragmatically odd, although not strictly speaking impossible. For a somewhat more extensive discussion of the role of perspective in (some cases of) polarity licensing, I refer to Hoeksema (2021).

Half a brain has the curious property of being a polarity item that is restricted to conditional clauses, and relative clauses restricting universal or negative quantifiers. Here too, questions spring up. Are there more expressions like it, and are they polarity sensitive? If there are, and they are not NPIs, then my suggestion that the unusual distribution does not have to be treated in terms of polarity licensing might be on the right track. In any case, we need to think more on such cases and be on the lookout for them.

In addition, we need to think more about *any* in predicate nominals. We discussed *if history is any guide*. Why is it often so bad and only sometimes any good? Think about pairs such as

- (43) A teacher should not hit/*be any student.
- (44) She is not a/*any girl anymore.
- (45) Don’t be a/*any stranger!

But also think about:

- (46) Hotze is not just any professor.
(acceptable, thanks to the *just*)
- (47) Mr. Chairman, the letter bears internal evidence that Mr. Newbold is not any friend of mine. [US Senate committee hearings, 1923]
(*is not any friend of mine* sounds better than *is not any friend*)
- (48) Jones is not anyone's enemy.
(better than *Jones is not any enemy.*)

Sentences such as *history is any guide* only appear in conditional clauses. They are stereotypical but not fully ossified. How best to treat them? And are there more of them?

As a beginning linguist, I sometimes wondered how complex linguistics really is. Some areas are pretty much finished. The phonemic inventory of standard English is not something you would want your students to write a dissertation about. Polarity items ditto, I thought — how hard can they be? Forty years later, I am still trying to solve parts of the puzzle. Dear Hotze, I fear our work may never be done.

References

- Austin, John L. 1956. Ifs and cans. *Proceedings of the British Academy* 42:109–132.
- Cambridge Dictionary. 2023. URL: <https://dictionary.cambridge.org/dictionary/english/i-ll-be-hanged-if-i-know> [last accessed on 20 July 2023].
- Crnič, Luka. 2019. *Any*: Logic, likelihood, and context. *Language and Linguistics Compass* 13(11):1–20.
- Csipak, Eva. 2014. Minimizers in conditional threats and promises. *Proceedings of Sinn und Bedeutung* 18:95–109.
- Gajewski, Jon. 2011. Licensing strong NPIs. *Natural Language Semantics* 19(2):109–148.
- Giannakidou, Anastasia. 2011. Negative and positive polarity items: licensing, compositionality and variation. In Klaus von Stechow, Claudia Maienborn, & Paul Portner (eds.) *Semantics: An*

international handbook of natural language meaning. Berlin, Germany: Mouton de Gruyter, 1660–1712.

Guerzoni, Elena. 2003. *Why even ask? On the pragmatics of questions and the semantics of answers*. Doctoral dissertation, MIT.

Heim, Irene. 1984. A note on negative polarity and downward entailingness. *Proceedings of NELS* 14:98–107.

Hoeksema, Jack. 2012. On the natural history of negative polarity items. *Linguistic Analysis* 38(1):3–33.

Hoeksema, Jack. 2021. Verbs of deception, point of view and polarity. *Proceedings of the 28th International Conference on Head-Driven Phrase Structure Grammar*. Frankfurt/Main, Germany: University Library, 26–46.

Hoeksema, Jack. 2022. Conditional interpretation of *with/without* and the licensing of negative polarity items. In Hedwig Sekeres & Martijn Wieling (eds.), “*ôfskie / ôfschaaid*”, *liber amicorum Goffe Jensma*. Groningen, Netherlands, 31–38.

Horn, Laurence R. 1989. *A Natural History of Negation*. Chicago, IL: University of Chicago Press.

Klima, Edward. 1964. Negation in English. In Jerry A. Fodor & Jerrold J. Katz (eds.), *The Structure of Language. Readings in the Philosophy of Language*. Hoboken, NJ: Prentice Hall, 246–323.

Lakoff, Robin. 1969. Some reasons why there can't be any *some-any* rule. *Language* 45:608–615.

Lee, Young-Suk & Laurence Horn. 1994. *Any as indefinite plus even*. Manuscript, Yale University.

Rawlins, Kyle. 2020. Biscuit conditionals. In Daniel Gutzmann, Lisa Matthewson, Cécile Meier, Hotze Rullmann, & Thomas Ede Zimmermann (eds.), *The Wiley Blackwell companion to semantics*. Hoboken, NJ: John Wiley & Sons, 1–26.

- Reuneker, Alex. 2016. Conditional use of prepositional phrases in Dutch: The case of *zonder* (“without”). In Jenny Audring & Sander Lestrade (eds.), *Linguistics in the Netherlands 2016*. Amsterdam, Netherlands: John Benjamins, 121–134.
- Rullmann, Hotze. 1996. Two types of negative polarity items. *Proceedings of the North East Linguistic Society* 26:335–350.
- Rullmann, Hotze. 2003. Additive particles and polarity. *Journal of Semantics* 20:329–401.
- von Stechow, Kai. 1999. NPI licensing, Strawson entailment, and context dependency. *Journal of Semantics* 16(2):97–148.
- Wouden, Ton van der. 1994. *Negative contexts*. Doctoral dissertation. University of Groningen.
- Wouden, Ton van der. 2001. Three modal verbs. In Sheila Watts, Jonathan West, & Hans-Joachim Solms (eds.), *Zur Verbmorphologie germanischer Sprachen*. Tübingen, Germany: Niemeyer, 189–210.
- Zwarts, Frans. 1981. Negatief polaire uitdrukkingen I. *GLoT* 4(1):35–132.
- Zwarts, Frans. 1998. Three types of polarity. In Fritz Hamn & Erhard Hinrichs (eds.), *Plurality and Quantification*. Dordrecht, Netherlands: Springer, 177–238.

Variable strength in necessity modality: A case of variation between Afrikaans, Dutch, English, and German

VERA HOHAUS, SIENA WEINGARTZ, & MARIJN BOOMARS
The University of Manchester

1 Introduction

In their landmark paper, Rullmann et al. (2008) discuss the perceived variability in the quantificational force of modal expressions in St'át'imcets (Lillooet Salish), which they then derive from an underlyingly uniform semantics. The paper has triggered a productive line of research into such (often only prima-facie) variability in modal meaning across languages (Deal 2011; Bochnak 2015; Jeretič 2021; Newkirk 2022; among others).

We present here a related case of variability, in the strength of some of the cognates of the English necessity modal *must* in three other Germanic languages. In English, a distinction in the strength of a necessity can be observed between *must* and the weaker *ought to* in the infamous (1), for instance. The distinction has been argued to be lexically encoded (see, in particular, von Stechow & Iatridou 2008, 2023; Rubinsteyn 2012). While German *müssen*, in (2), appears to encode strong necessity just like its English counterpart, the celebrant's native Dutch *moeten* as well as Afrikaans *moet*, in (3) and (4), systematically allow for weak and strong readings, even in the absence of consequent x-marking, which is required in German (see also Matthewson & Truckenbrodt 2018).

(1) ENGLISH:

*Employees **must** wash their hands.*
*Non-employees really **ought to** wash their hands, too.*

(2) GERMAN:

*Alle Mitarbeiter **müssen** sich die Hände waschen.*
all employees NEC REFL the hands wash
*Alle anderen **sollten** das eigentlich auch.*
all others NEC+X that EIGTL also

(3) DUTCH:

Werknemers moeten hun handen wassen.

employees NEC their hands wash

Niet-werknemers moeten eigenlijk ook hun handen wassen.

non-employees NEC EIGTL also their hands wash

(4) AFRIKAANS:

Werkers moet hande was.

workers NEC hands wash

Nie-werkers moet eintlik ook hulle hande was.

not-workers NEC EIGTL also their hands wash

Building on Weingartz & Hohaus (to appear), we suggest that Afrikaans *moet* and Dutch *moeten* are underlyingly weak, unlike their English and German cognates. Within a domain-restriction approach to weak necessity (von Fintel & Iatridou 2008; Rubinstein 2012; Vander Klok & Hohaus 2020), this proposal can be formalised as follows: Afrikaans *moet* and Dutch *moeten* lexically specify for a secondary ordering source, unlike English *must* and German *müssen*, but allow for this secondary ordering source to be empty, unlike English *ought to* and *should*, for instance.

Our focus in this short paper is to present an initial data set that supports such an analysis and invites further research into the variability in strength and the variation between the four languages. We set the discussion of the data against a short background section, Section 2, which introduces some defining characteristics of weak necessity modal expressions and surveys the morpho-syntactic strategies that languages adopt for weak strength. Section 3 presents the data. Section 4 sketches the analysis and directions for future research.

2 Background

As a third descriptive dimension of modal meaning, we take strength here not to be a distinction in force (that is, the distinction between possibility and necessity), nor in flavour (that is, broadly, the distinction between epistemic and root). Weak necessity (WN, henceforth) modal expressions behave like their strong counterparts, and unlike possibility modal expressions, in that they do not allow for the conjunction of mutually exclusive propositions (Rubinstein 2012, 2021), as shown in

(5). They entail possibility, while being entailed by the stronger counterpart (Horn 1972; von Fintel & Iatridou 2008; Rubinstein 2021). This entailment relation gives rise to a scalar implicature to the exclusion of the stronger necessity. The implicature can be overtly reinforced, cancelled, or suspended, as shown in (6). With Rubinstein (2021), we take these two properties to be defining properties of weak necessity, which allows us to identify it across languages.

(5) Coordination:

- a. #*You **must** / **should** stay but you also **must** / **should** go.*
- b. *You **can** stay but you **can** also go.*

(6) Scalar implicature:

*I **ought to** help the poor.*

- a. Reinforcement: *But I don't have to.*
- b. Cancellation: *In fact, I must.*
- c. Suspension: *Maybe I have to.*

We additionally discuss below a third property of weak modal expressions, relating to their acceptability in contexts in which there are no feasible alternatives (Sloman 1970; Sæbø 2001; von Fintel & Iatridou 2008), illustrated in (7), which is inspired by von Fintel and Iatridou (2008:118).¹

(7) Lack of alternatives:

Context: The only way to get to Harlem by train is the A-line.

- a. #*If you want to go to Harlem by train, you **should** take the A-line.*
- b. *If you want to go to Harlem by train, you **must** take the A-line.*

We frame the formal discussion within the domain-restriction approach to weak necessity (von Fintel & Iatridou 2008; Rubinstein 2012; but see

¹ We however set aside two further properties that have featured in the literature on weak necessity in English, namely, the negotiability of the additional considerations that inform the modal claim (Rubinstein 2012, 2021) and the gradability of these modals (Klecha 2014; Portner & Rubinstein 2016; among others).

Portner & Rubinstein 2016; Agha & Jeretič 2022; inter alia), under which weak strength is a result of a smaller domain of quantification: “Strong necessity modals say that the prejacent is true in all of the favoured worlds, while weak necessity modals say that the prejacent is true in all of the very best (by some additional measure) among the favoured worlds” (von Fintel & Iatridou 2008:118). One possible implementation of such an approach is sketched in (8) to (10). A strong necessity modal like English *must* under this analysis quantifies over the best of the ordered accessible worlds. The domain of quantification for a weak necessity modal like English *should* are the best of those best worlds, ordered with respect to the propositions provided by the secondary ordering source, and hence a subset of the domain of quantification of the strong counterpart.

- (8) For any accessibility relation $a \in D_{\langle s, \langle s, t \rangle \rangle}$, ordering source $o \in D_{\langle s, \langle \langle s, t \rangle, t \rangle \rangle}$, proposition $p \in D_{\langle s, t \rangle}$, and possible world $w \in D_s$,
 \llbracket (strong necessity) $\rrbracket(a)(o)(p)(w) = 1$ iff
 $\forall w' \in \text{BEST}_{(a(w), o(w))}: p(w') = 1$
- (9) For any accessibility relation $a \in D_{\langle s, \langle s, t \rangle \rangle}$, primary ordering source $o1 \in D_{\langle s, \langle \langle s, t \rangle, t \rangle \rangle}$, secondary ordering source $o2 \in D_{\langle s, \langle \langle \langle s, t \rangle, t \rangle, t \rangle \rangle}$, proposition $p \in D_{\langle s, t \rangle}$, and possible world $w \in D_s$,
 \llbracket (weak necessity) $\rrbracket(a)(o1)(o2)(p)(w) = 1$ iff
 $\forall w' \in \text{BEST}(\text{BEST}_{(a(w), o1(w))}, o2) : p(w') = 1$
- (10) a. For any set of worlds $W \in D_{\langle s, t \rangle}$ and set of propositions $P \in D_{\langle \langle s, t \rangle, t \rangle}$: $\text{BEST}(P, W) = \{w \in D_s : \neg \exists w' \in W : w' >_P w\}$
 b. For any set of worlds W and set of propositions $P : \forall w, w' \in W : w >_P w'$ iff $\{p \in P : p(w') = 1\} \subset \{p' \in P : p'(w) = 1\}$
 (see also von Fintel & Heim 2011:61, no. 107)

Languages may lexicalise the above distinction, like English, but may also adopt a morphologically more transparent route: Under this morpho-syntactic strategy, weak necessity is marked on a strong expression, recruiting the verbal morphology that also appears in the consequent of counterfactual conditionals (= consequent x-marking, von Fintel & Iatridou 2008, 2023), or specialised morphology (Vander Klok & Hohaus 2020). Languages may additionally resort to comparative paraphrases to convey weak necessity (Rubinstein 2014). Weingartz and Hohaus (to appear) discuss a fourth, previously unattested strategy: In

Afrikaans and Samoan (Austronesian, Oceanic), the distinction between weak and strong necessity may be left unmarked. Note that these strategies are not mutually exclusive within a language: Dutch, for instance, makes use of all four. We illustrate the lexical and morphological strategy in (11) and (12), respectively (the latter from von Fintel & Iatridou 2008:124). An example of a relevant comparative is in (13).²

(11) DUTCH:

Lexicalised weak necessity with *horen* ‘to befit’:

Je hoort eigenlijk een elektrische toets te doen.
 you WNEC EIGTL a electrical test to do
Om je de waarheid te zeggen, je moet dat doen.
 to you the truth to say you NEC that do
 ‘You should actually do an electrical test. To tell you the truth,
 you have to.’

(12) DUTCH:

Consequent x-marking with *zou, zouden* ‘would’:

a. *Als ik rijk was, zou ik stoppen met werken.*
 if I rich were X I stop with work
 ‘If I were rich, I would stop working.’

b. *Je zou eens Anna Karenina moeten lezen, maar*
 you X sometime NAME NEC read but
het hoeft niet.
 it NEC not
 ‘You should read *Anna Karenina* sometime, but you don’t
 have to.’

(13) DUTCH:

Comparative paraphrase:

Het is beter dat je gaat.
 it is better that you go
 ‘You better go.’

² Abbreviations used in glosses include EIGTL = cognates of German *eigentlich* ‘actually, technically’, NEC = necessity, POS = possibility, REFL = reflexive, WNEC = weak necessity, and X = consequent x-marking.

The unmarked case is the topic of the next section, where our focus is on the strength of English *must* and its cognates in Afrikaans, Dutch, and German. Building on Weingartz and Hohaus (to appear), we highlight an interesting point of variation between these closely related languages: Dutch patterns with Afrikaans rather than English and German, in that it allows weak modal strength to go morphologically unmarked.

3 The case for variable strength

The cognates of English *must* in all three languages are standardly described as necessity modal expressions (Kratzer 1978 et seqq; Zifonun et al. 1997; Diewald 1999; Matthewson & Truckenbrodt 2018; de Villiers 1971; Donaldson 1993; Haeseryn et al. 1997; Huitink 2012; to name but a few for each language). They also pattern as such with respect to the coordination diagnostic for necessity from the previous section, as shown in (14).

(14) Coordination:

a. ENGLISH:

*#Dogs **must** stay outside, and they **must** stay inside.*

b. GERMAN:

*#Bei uns **müssen** Hunde draußen bleiben,
at us NEC dogs outside stay
und sie **müssen** im Haus bleiben.
and they NEC in+the house stay*

c. AFRIKAANS:

*#Die hond **moet** buite bly en hy **moet** ook inkom.
the dog NEC outside stay and he NEC also come.in*

d. DUTCH:

*#De hond **moet** buiten blijven en hij **moet** ook
the dog NEC outside stay and he must also
binnenkomen.
come.inside*

While we find Afrikaans *moet*, Dutch *moeten*, and German *müssen* in contexts that target a strong interpretation, like (15), Afrikaans *moet* and Dutch *moeten* are also attested with weak interpretations. Two naturally occurring examples are in (16) and (17). Note that both examples would receive a strong interpretation when translated with German *müssen*.

(15) *Context: According to the law, you have to be over 16 to buy hair dye. Kirri wants to buy some, but Peta informs her:*

a. AFRIKAANS:

Nee, jy kan nie, want jy moet 16 of ouer wees.
 no you POS not because you NEC 16 or older be
 ‘No, you can’t, because you have to be 16 or older.’

b. DUTCH:

Nee, dat kan niet, want je moet 16 jaar of ouder zijn.
 no that POS not because you NEC 16 year or older
 zijn.
 be

c. GERMAN:

Nein, das geht nicht. Dafür musst du über 16 Jahren alt sein.
 no that goes not there.for NEC you over 16 years
 alt sein.
 old be

(16) AFRIKAANS:

Context: From a horoscope.

Jy moet nou ekstra versoorg tref.
 you NEC now extra precaution meet
 ‘You should now take extra precaution.’ (Donaldson 2002:47)

(17) DUTCH:

Context: An IT consultancy on the recommended frequency of doing a cyber security risk assessments for your company.

*Hoe vaak moet ik zo'n assessment eigenlijk doen? Wat is
how often NEC I so+a assessment EIGTL do what is
de optimale frequentie?
the optimal frequency*

'How often should I do such an assessment? What is the optimal frequency?'³

Note that in German, *müssen* does not generate a scalar implicature to the exclusion of a stronger necessity, and the continuations in (18) are contradictory. Compare this to Afrikaans (19), however, where the weak interpretation of *moet* seems to generate an implicature that can be targeted in the continuation both by strong necessity (and strong negative polarity) *hoef* 'to have to' and a strong interpretation of *moet*. In Dutch, in (20), evidence in favour of such an implicature comes from reinforcement with *hoeven* in (20a), but (20b) and (20c) are judged as unacceptable. These data may suggest a preference for speakers of Dutch to keep the strength of *moeten* constant within a sentence, but warrants further investigation.⁴

(18) GERMAN:

*Ich muss nachher noch beim Sommerfest vorbeischauen.
I NEC later still at+the summer.party look.by
'I still need to go the summer party.'*

³ QS Solutions, "Een security-assessment, hoe vaak moet je dat eigenlijk doen?" (URL: <<https://qssolutions.nl/blogs/een-security-assessment-hoe-vaak-moet-je-dat-eigenlijk-doen/>>, last accessed 24th January 2024).

⁴ Tine Breban (p.c.) suggests (i) below as an acceptable continuation, instead of (20b), which has the desirable interpretative effect but does not rely on a repetition of *moeten*.

(i) Dutch:

*Strikt genomen is het niet absoluut noodzakelijk dat ik ga.
strictly taken is it not absolutely necessary that I go
'In fact, it is not absolutely necessary that I go.'*

- a. Reinforcement: #*Aber **müssen** tue ich das natürlich*
 but NEC do I that naturally
nicht.
 not
 ‘But of course I don’t have to.’
- b. Cancellation: #*Genaugenommen **muss** ich dahin.*
 strictly.taken NEC I thither
 ‘In fact, I have to.’
- c. Suspension: #*Vielleicht **muss** ich das sogar.*
 maybe NEC I that even
 ‘Maybe I even must.’

(19) AFRIKAANS:

*Ek **moet** nog na die partyjie toe gaan!*

I NEC still to the party to go

‘I should still go to the party.’

- a. Reinforcement: *Maar eintlik **hoef** ek nie te gaan nie.*
 But EIGTL NEC I not to go not
 ‘But actually, I don’t have to.’
- b. Cancellation: *Streng gesproke, **moet** ek gaan.*
 strictly spoken NEC I go
 ‘Strictly speaking, I have to go.’
- c. Suspension: *Miskien **moet** ek maar gaan.*
 maybe NEC I but go
 ‘Perhaps, I must.’

(20) DUTCH:

*Ik **moet** later nog naar het zomerfeest.*

I NEC later still to the summer.party

‘I still need to go to the summer party later.’

- a. Reinforcement: *Maar ik **hoef** dat niet.*
 but I NEC that not
 ‘But I don’t have to.’

- b. Cancellation: #*Strikt genomen moet ik daarheen.*
 strictly taken NEC I there.to
 ‘Strictly speaking, I must.’
- c. Suspension: #*Misschien moet ik dat zelfs.*
 maybe NEC I that even
 ‘Maybe I even must.’

The last data set suggestive of weak strength for Afrikaans *moet* and Dutch *moeten* is in (21) and (22). German *müssen* patterns with English *must* when it comes to its acceptability in contexts that establish several alternatives, like (21). Afrikaans *moet* and Dutch *moeten* are judged acceptable (and thus behave like weak necessity *should* in English and German x-marked *sollen* in this context). They are however also acceptable in (22), which targets a strong necessity reading.

(21) *Context: There are three ways to get to Manchester: The back routes, the M6, and through Reading. Bess says that the route using the M6 is best. So, according to her:*

a. GERMAN:

#*Nach Manchester **musst** du die Autobahn nehmen.*
 to NAME NEC you the motorway take
 Lit. ‘To Manchester, you must take the motorway.’

b. AFRIKAANS:

*As jy na Manchester toe gaan, **moet** jy die M6*
 if you to NAME to go NEC you the NAME
gebruik.
 use

c. DUTCH:

*Als je naar Manchester gaat, **moet** je de M6*
 if you to NAME go NEC you the NAME
nemen.
 take

(22) Lack of Alternatives:

Context: There are usually three ways to Manchester: The back routes, the M6, and through Reading. Currently the M6 is the only option; the other roads are closed.

a. GERMAN:

*Nach Manchester **musst** du im Moment die*
to NAME NEC you at+the moment the
Autobahn nehmen.
motorway take

‘To Manchester, at the moment, you must take the motorway.’

b. AFRIKAANS:

*Jy **moet** die M6 gebruik, omdat die ander paaie*
you NEC the NAME use because the other roads
toe is.
closed are

c. DUTCH:

*Je **moet** de M6 nemen, omdat de andere*
you NEC the NAME take because the other
wegen dicht zijn.
ways closed are

We conclude from this brief discussion that Afrikaans *moet* and Dutch *moeten* appear to exhibit variable strength, unlike their English and German cognates.

4 Discussion

To our knowledge, this variability is a previously largely unexplored point of variation within the Germanic languages, which also opens up interesting perspectives for further synchronic and diachronic research, especially in the light of the variability in force discussed in Yanovich (2016) for Old English *motan* and Middle English *moten*. From the perspective of the crosslinguistic typology of the dimension of modal meaning then, not only are flavour and force subject to variability across languages, but strength is as well.

Weingartz (2022) and Weingartz and Hohaus (to appear) suggest that such variability in strength can be systematically derived from a uniform semantics if we assume that some weak necessity expressions allow for an empty secondary ordering source. Under an empty secondary ordering source, the weak necessity claim ends up equivalent to a strong necessity claim with a single ordering source (see also Rubinstein 2013). English *should* and *ought* under such a view would lexically prohibit an empty secondary ordering source. Afrikaans *moet* and, as we tentatively propose here, Dutch *moeten* would lexically specify for a secondary ordering source, as sketched in (9) above, but would also allow for it to be empty, depending on context.

In addition to context, other lexical material may possibly also interact with the strength of a modal expression. The observant reader will have noticed that the weak interpretations of Afrikaans *moet* and Dutch *moeten* co-occur with the discourse particle *eintlik* in Afrikaans, or *eigenlijk* in Dutch, in many of the above examples. While not obligatory throughout, there is a strong preference for its use with weak interpretations, although, in German, the use of *eigentlich* is not enough to bring about a weak interpretation of *müssen*. We will leave the exploration and analysis of this interaction for another occasion, or for the enjoyment of the celebrant. *Proficiat, Hotze!*

References

- Agha, Omar & Paloma Jeretič. 2022. Weak necessity modals as homogeneous pluralities of worlds. *Proceedings of Semantics and Linguistic Theory (SALT)* 32:831–851.
- Bochnak, M. Ryan. 2015. Variable force modality in Washo. *Proceedings of the Annual Meeting of the North Eastern Linguistics Society (NELS)* 45:105–114.
- Deal, Amy R. 2011. Modals without scales. *Language* 87(3):559–585.
- Diewald, Gabriele. 1999. *Die Modalverben im Deutschen*. Tübingen, Germany: Niemeyer.
- Donaldson, Bruce C. 2002. *Colloquial Afrikaans*. London, UK: Routledge.
- Donaldson, Bruce C. 1993. *A grammar of Afrikaans*. Berlin, Germany: De Gruyter.

- von Fintel, Kai & Irena Heim. 2011. *Intensional semantics*. Lecture notes, Massachusetts Institute of Technology, Cambridge, MA.
- von Fintel, Kai & Sabine Iatridou. 2023. Prolegomena to a theory of x-marking. *Linguistics and Philosophy* 46(6):1467–1510.
- von Fintel, Kai & Sabine Iatridou. 2008. How to say OUGHT in Foreign: The composition of weak necessity modals. In Jacqueline Guéron & Jacqueline Lecarme (eds.), *Time and modality*. Dordrecht, Netherlands: Springer, 115–141.
- Haeseryn, Walter. 1997. *Algemene Nederlandse spraakkunst* (2nd ed.). Groningen, Netherlands: Nijhoff.
- Herburger, Elena & Aynat Rubinstein. 2019. Gradable possibility and epistemic comparison. *Journal of Semantics* 36(1):165–191.
- Horn, Laurence R. 1972. *On the semantic properties of logical operators in English*. Doctoral dissertation, University of California, Los Angeles, CA.
- Huitink, Janneke. 2012. Modal concord: A case study of Dutch. *Journal of Semantics* 29(3):403–437.
- Klecha, Peter. 2014. *Bridging the divide: Scalarity and modality*. Doctoral dissertation, University of Chicago, Chicago, IL.
- Jeretič, Paloma. 2021. *Neg-Raising modals and scaleless implicatures*. Doctoral dissertation, New York University, NY.
- Kratzer, Angelika. 1978. *Semantik der Rede: Kontexttheorie, Modalwörter, Konditionalsätze*. Königstein, Austria: Scriptor.
- Matthewson, Lisa & Hubert Truckenbrodt. 2018. Modal flavour/ modal force interactions in German. *Linguistische Berichte* 256:259–313.
- Newkirk, Lydia 2022. *Be flexible, but not too flexible: Variable-force modals in Kinande and the typology of modal force*. Doctoral dissertation, Rutgers, The State University of New Jersey, New Jersey, NJ.
- Portner, Paul & Aynat Rubinstein. 2016. Extreme and non-extreme deontic modals. In Nate Charlow & Matthew Chrisman (eds.), *Deontic modality*. Oxford, UK: Oxford University Press, 256–282.
- Rubinstein, Aynat. 2021. Weak necessity. In Daniel Gutzmann, Lisa Matthewson, Cécile Meier, Hotze Rullmann, & Thomas Ede

- Zimmermann (eds.), *The Wiley Blackwell companion to semantics*, Vol 4. Hoboken, NJ: Wiley-Blackwell, 1–44.
- Rubinstein, Aynat. 2014. On necessity and comparison. *Pacific Philosophical Quarterly* 95(4):512–554.
- Rubinstein, Aynat. 2013. Figuring out what we ought to do: The challenge of delineating priorities. *Proceedings of the Annual Penn Linguistics Colloquium* 36:169–178.
- Rubinstein, Aynat. 2012. *Roots of modality*. Doctoral dissertation, University of Massachusetts, Amherst, MA.
- Rullmann, Hotze, Lisa Matthewson, & Henry Davis. 2008. Modals as distributive indefinites. *Natural Language Semantics* 16(4):317–357.
- Sæbø, Kjell J. 2001. Necessary conditions in natural language. In Caroline Féry & Wolfgang Sternefeld (eds.), *Auditor Vox Sapientia: A Festschrift for Arnim von Stechow*. Berlin, Germany: Akademie Verlag, 427–449.
- Slooman, Aaron. 1970. *Ought and better*. *Mind* 79(315):385–394.
- Vander Klok, Jozina & Vera Hohaus. 2020. Weak necessity without weak possibility: The composition of modal strength distinctions in Javanese. *Semantics and Pragmatics* 13(12):1–50.
- de Villiers, Meyer. 1971. *Die grammatika van tyd en modaliteit*. Cape Town, South Africa: Balkema.
- Weingartz, Siena. 2022. *Modal necessity in Ndebele (Bantu; Zimbabwe) and Afrikaans (Germanic; South Africa)*. M.A. dissertation, The University of Manchester, Manchester, UK.
- Weingartz, Siena & Vera Hohaus. (to appear). Variable modal strength in Afrikaans and Samoan: Deriving strong necessity from weak necessity. *Proceedings of TripleA* 10.
- Yanovich, Igor. 2016. Old English **motan*, variable-force modality, and the presupposition of inevitable actualisation. *Language* 92(3):489–521.
- Zifonun, Gisela, Ludger Hoffmann, & Bruno Strecker. 1997. *Grammatik der deutschen Sprache*. Berlin, Germany: De Gruyter.

On the absence of evidence for nominal tense: The ʔayʔajuθəm past marker *-ot**

MARIANNE HUIJSMANS
University of Alberta

1 Introduction

In a number of Salish languages, the same morpheme that is used to indicate the past on verbs also occurs on nouns (e.g., Burton 1997, Wiltschko 2003, Matthewson 2005). On nouns, the past marker typically indicates that the referent of the DP is dead or destroyed, and for possessed nouns, the past marker can indicate that the possession relation no longer holds. The past marker on nouns is therefore frequently translated into English using adjectives like ‘late’, ‘former’, or ‘ex’.

Prior analyses differ in whether they treat the past marker as actually marking nominal tense. Burton (1997) proposes that the past marker on nouns in Halkomelem encodes past tense on nouns, while Wiltschko (2003) argues that it realizes an interpretable tense feature on D. In contrast, Matthewson (2005) argues that the ‘past tense’ morphemes found on nouns in St’át’imcets and Halkomelem are really temporal modifiers that

* I am deeply grateful to all the speakers who have shared their language and insights with me over the years, especially Elsie Paul, Phyllis Dominic, Freddie Louie, Betty Wilson, and late Joanne Francis, whom I consulted at different points on this topic. Without their resilience and dedication to keeping their language alive, this work would not be possible. I am also grateful to the audience at ICSNL 54 for helpful discussion of a much earlier version of part of this work. Finally, I would like to acknowledge the Jacobs Research Funds, a SSHRC Insight grant #435-2016-1694 awarded to Henry Davis, and start up funds from the University of Alberta that have made this research possible.

This paper is dedicated to Hotze Rullmann, who has been an important mentor throughout my development as a semanticist. Though Hotze was not officially on my committee during my time as a graduate student at UBC, he has always been very generous with feedback on presentations, abstracts, and papers, and his feedback has never failed to be important, both for his attention to detail and his precision of thought. I have also learnt a lot from him through participating in several research groups with him, collaborating with him on several occasions, and being his teaching assistant for an upper-level undergraduate semantics class. Thank you, Hotze, and wishing you a very happy birthday!

optionally attach to nouns and verbs. Under this analysis, tense is not part of the functional architecture of DPs in St'át'imcets and Halkomelem.

A similar debate exists for past markers in Guaraní languages. Thomas (2012, 2014) argues that a past marker that occurs on nouns in Mbyá is a nominal tense, while Tonhauser (2006, 2007) argues that its cognate in Paraguayan Guaraní is a predicate modifier that is not a true nominal tense.

In this paper, I examine the cross-category use of the past marker in another Salish language, ʔayʔajuθəm (a.k.a. Comox-Sliammon; ISO: coo). I argue that the past marker *-oł* in ʔayʔajuθəm does not occupy T in either clausal or nominal contexts, presenting novel data that shows that *-oł* does not have a fixed syntactic position and can apply to different constituents. I therefore argue that in both clausal and nominal environments, *-oł* acts as a temporal modifier and provide a preliminary semantics where it combines with a predicate to add a presupposition restricting the reference time for the predicate to the past.

My analysis of *-oł* therefore supports Matthewson's (2005) position that the past markers in Halkomelem and St'át'imcets are temporal modifiers rather than tense. Crucially, following Matthewson's argumentation for these other languages, since *-oł* is not a morphological realization of T in ʔayʔajuθəm, its presence on nouns cannot be used to argue for a tense projection in nominal environments or a tense feature on D.

This paper is organized as follows: In Section 2, I provide language background, and in Section 3, I provide theoretical background. In Section 4, I briefly discuss ʔayʔajuθəm tense and the temporal interpretation of DPs. In Section 5, I discuss the interpretation of *-oł* on verbs, nouns, and adjectives. In Section 6, I present arguments that *-oł* is not a past tense but rather a temporal modifier. In Section 7, I provide a preliminary analysis of *-oł*.

2 Language background

ʔayʔajuθəm is a Central Salish language, the ancestral language of the Tla'amin, Homalco, Klahoose, and K'ómoks Nations,¹ whose traditional territory lies along the northern Georgia Strait. Due to the impacts of colonialism, especially the residential school system (TRC 2015), only 3% of the traditionally ʔayʔajuθəm-speaking population are now first-language speakers, while 10% are second-language learners (FPCC 2022). There

¹ Pentlatch and Kwak'wala are also ancestral languages of the K'ómoks Nation.

is currently an active and determined push for reclamation of language and culture among the four nations.

I consulted with five Elders from Tla'amin, Klahoose, and Homalco at different points during the background research for this paper but have worked especially closely with one speaker from Tla'amin in the later stages of this research.

3 Theoretical background

I follow much previous literature (Klein 1994, et seq.) in assuming that tense provides the evaluation time for a proposition, known as the *reference time* (RT). Tense relates this RT to a temporal anchor. In matrix clauses, the temporal anchor is typically the utterance time (UT). If the tense is present, the RT for the proposition is the same as the temporal anchor, while if the tense is past, the RT for the proposition precedes the temporal anchor.

- (1) a. The sky is blue. (RT = UT)
 b. The sky was blue. (RT < UT)

Enç (1981, 1986) points out that the temporal interpretation of DPs is at least partly independent of the temporal interpretation of the clause they appear in. A classic example is given in (2):

- (2) Every fugitive is now in jail. (Enç 1986:409)

Although the sentence has present tense, the sentence is not about individuals who are fugitives now (or it would be contradictory), but rather about individuals who were fugitives before but are now in jail. To capture this, Enç proposes that each noun must have its own temporal argument or *NP evaluation time*. However, the temporal argument of a noun need not be syntactically represented as a nominal tense but may be rather supplied by the context (Enç 1986:422).

The question that this paper aims to address is whether the presence of the past marker on nouns in ʔayʔajuθəm should be taken as evidence that ʔayʔajuθəm nouns contain tense in their *syntax*.²

² Burton's (1997) analysis of the past marker on nouns in Halkomelem as the morphological realization of nominal tense seems to imply a tense projection, although he does not explicitly claim a T projection in nominal environments.

4 ʔayʔajuθəm tense

Following Matthewson (2006) for St’át’imcets, I analyze ʔayʔajuθəm as having a null nonfuture tense (see Huijsmans 2022). Unmarked predicates may be interpreted as holding of a past or present time, as illustrated in (3), depending on the context (and subject to aspectual restrictions that will not concern us here; see Huijsmans 2022:30–34 for discussion).³

- (3) a. *Context: Talking about a cat in the room.*
 kʷət gi Patlik. ʔeʔɛltən.
 kʷə[n]-t=gi Patlik. ʔi~ʔiltən.
 see-CTR=DPRT Patrick PROG~eat
 ‘Look at Patrick. He’s eating.’ [PRESENT] (vf | JF.2018/05/01)
- b. tʰətʰɣʷtəs tə cars skʷijol ʔi
 tʰə~tʰɣʷ-t-as tə=car-s skʷijul ʔiy
 PROG~wash-CTR-3ERG DET=car-3POSS this.morning CONJ
 həwtəm sçitʔos.
 hiwt-əm s=çəl-ʔu+s
 ahead-MD NMLZ=rain-PST-3POSS
 ‘He was washing his car this morning before it rained. [PAST]
 (vf | PD.2019/04/10)

The future is obligatorily marked with the future clitic *səm* (Huijsmans and Mellesmoen 2021, Huijsmans 2022:28). In (4), for example, it is not possible to interpret the cooking event in the future of the UT without the future clitic.

³ The abbreviations used in this paper follow Leipzig glossing conventions with some additions: ACT.INTR ‘active intransitive’, CLD ‘clausal demonstrative’, CLF.PRT ‘clefing particle’, CTR ‘control transitivizer’, DPRT ‘discourse particle’, EPEN ‘epenthetic segment’, INFER ‘inferential’, INT ‘intensifier’, MD ‘middle’, NCTR ‘non-control transitivizer’, RPT ‘reportative’, SBRD ‘subordinate’. The top line of each examples is an orthographic representation, and the second line is a roughly phonemic representation using NAPA. ‘vf’ stands for volunteered form, a form provided by the speaker, while ‘sf’ stands for suggested form, a form suggested to the speaker by the researcher.

- (4) *Context: I'm making a plan for dinner since someone gave me a fish.*

a. #č̣εχatč̣ tə ʔɛnx^w snanat.
 č̣əǰ-at=č̣ tə=ʔanx^w s=nanat
 cook-CTR=1SG.SBJ DET=fish NMLZ=evening
 ‘I’ll cook the fish tonight.’

b. č̣εχatt^θəḿ tə ʔɛnx^w snanat.
 č̣əǰ-at=t^θ+əḿ tə=ʔanx^w s=nanat
 cook-CTR=1SG.SBJ+FUT DET=fish NMLZ=evening
 ‘I’ll cook the fish tonight.’

(Huijsmans and Mellesmoen 2021:106)

Formally, I analyze the null non-future tense as in (5) (Huijsmans 2022:28, originally from Matthewson 2006:680 for St’át’imcets). It is a pronominal tense, bearing an index *i* interpreted by the assignment function *g*. It is restricted to non-future times by a presupposition that no part of the RT interval $g(i)$ follows the UT t_0 .

- (5) $\llbracket \text{NON-FUT}_i \rrbracket^{\text{g,c}} = g(i)$; defined only if no part of $g(i)$ is after t_0

As in English, the temporal interpretation of DPs in ʔayʔajuθəḿ is at least partly independent of the temporal interpretation of the clause as a whole. For instance, in (6), the RT for the clause is a past time when the speaker’s father was a child. However, the referent of the DP was not yet a father at that past reference time. The evaluation time for t^θ *man* ‘my father’ is rather the present.

- (6) *Context: I'm telling you about one of the neat things my dad did as a boy.*

hiyʔaʔmoł t^θ man k^w nənx^wiʔəḿ
 həy-ʔəḿ-uł t^θ=man k^w=nənx^wiʔəḿ
 make-ACT.INTR-PST 1SG.POSS=father DET=small.boat
 sčuy^os.
 s=čuy-ʔu+s
 NMLZ=child-PST+3POSS

‘My dad made a little boat when he was a kid.’

(sf | BW.2023/08/03)

5 Cross-category use of the ʔayʔajuθəm past marker

There is an optional past suffix *-ol*, which is the focus of this paper. It occurs on verbs, nouns, and adjectives, as shown in (7) (see Watanabe 2003:483–484 for a brief discussion of the use of *-ol* on nouns and verbs and Huijismans 2023 for arguments that nouns, verbs, and adjectives are syntactically and morphologically distinct categories). The adjective in (7c) is serving as the main predicate; predicative nouns and adjectives are not accompanied by a copula (see Huijismans 2023).

(7) a. kʷonetolč šε xʷipomixʷtən sʃεsol
 kʷən-í-t-ul=č šə=xʷip-umixʷ-tən sʃasul
 see-STAT-CTR-PST=1SG.SBJ DET=sweep-ground-INS yesterday
 nεʔ tə shed.
 niʔ tə=shed
 be.there DET=shed
 ‘I saw the broom yesterday in the shed.’
 [VERB] (vf | EP.2021/07/24)

b. čkʷa kʷonoxʷ šε totχʷlɑt
 čkʷa=kʷən-əxʷ šə=tutχʷlɑt
 1SG.SBJ=CLD=see-NCTR DET=necklace
 noʔos čε tʰ čičiyεʔol
 naʔ-ʔu+s=ča tʰ=č<ič>iya<ʔ>-ul
 OWN-PST+3POSS=INFER 1SG.POSS=grandmother<DIM>-PST
 χanatetol kʷa
 χan-at-it-ul=kʷa
 give-CTR-SBRD.PASS=RPT
 ʔə šεtʰ kʷukʷpaʔol.
 ʔə=šə=tʰ=kʷu<kʷ>pa<ʔ>-ul
 OBL=DET=grandfather<DIM>-PST
 ‘I found a necklace that must have belonged to my late grandmother that was given to her by my late grandfather.’ [NOUN]
 (vf | EP.2021/04/02)

c. pəqol tεʔε qəsnaɣ.
 pəq-ul tiʔi qəsnaɣ
 white-PST DEM shirt
 ‘This shirt used to be white.’ [ADJECTIVE] (sf | EP.2023/06/29)

In clausal contexts, the past marker is used to unambiguously establish a past RT. It therefore frequently shows up on verbs at the beginning of a narrative or discourse about a past time. For instance, in the *Hawaii Trip* storyboard (Underhill and Cable 2015), Bill answers Mary’s question about his summer. The speaker uses *-ol* in (8b) when Bill begins to talk about his trip, but not on the subsequent predicates in (8c–d).

- (8) a. haʔačx^w taʔačiš *this summer?*
 ha=a=čx^w taʔačiš *this summer*
 go=Q=2SG.SBJ travel *this summer*
 ‘Did you travel this summer?’
- b. hoʔoľč k^w *Hawaii*.
 hu-ʔuľ=č k^w=*Hawaii*
 go-PST=1SG.SBJ DET=*Hawaii*
 ‘I went to Hawaii.’
- c. ʔowuľč *plane eight* qəʃias k^wi.
 ʔuwuľ=č *plane eight* qəʃi=as k^wəy
 get.on=1SG.SBJ *plane eight still*=3SBJV *early*
 ‘I got on the plane at eight in the morning.’
- d. hoč təs k^w *Hawaii* k^w nat.
 hu=č təs k^w=*Hawaii* k^w=nat
 go=1SG.SBJ arrive DET=*Hawaii* DET=*night*
 ‘I arrived in Hawaii at night.’ (vf | PD.2019/04/10)

In nominal environments, that is, when the past occurs within an NP that is sister to a D,⁴ the past marker is typically used to indicate that the referent of the DP is dead, as in (7b), or destroyed, as in (9a); however, when the noun names a stage-level predicate, as in (9b), use of *-ol* can also indicate that the referent no longer has the nominal property, while still continuing to exist. On a possessed noun phrase, the interpretation can also be that the possession relation no longer holds (9c–d).

- (9) a. k^wak^wa qətx^w k^w ʔayeʔos.
 k^wa=k^wa=qətx^w k^w=ʔayaʔ-ʔu+s.
 RPT=CLD=burn DET=house-PST+3POSS
 ‘His has house burnt down (I heard).’ (vf | EP.2019/06/29)

⁴ The presence of D differentiates these environments from cases where the past marker occurs on nominal predicates, which are clausal environments.

- b. *Context: I talk to a former teacher at my highschool who is now retired. I never had him as a teacher myself. After, I tell my husband:*

kʷONOXʷolč šɛ tičɛhoł tuwa
 kʷən-əxʷ-ul=č šə=tiča-h-ul tuwa
 see-NCTR-PST=1SG.SBJ DET=teacher-EPEN-PST from

ʔətʰ kʷulawtxʷul.

ʔətʰ=kʷul-awtxʷ-ul

1SG.POSS=school-building-PST

‘I saw a former teacher from my school.’ (sf | EP.2021/11/20)

- c. *Context: Two friends are talking about a party this evening. They heard a mutual friend is coming. One of them realizes that their friend’s ex-wife may also come and says to her friend:*

čm sa ga qʷoləs šɛ saltuʔos?
 čəm=sa+ga qʷəl=as šə=saltu-ʔu+s
 what.is.with=FUT+DPRT come=3SBJV DET=wife-PST+3POSS

‘What if his ex-wife comes?’ (vf | EP.2021/05/21)

- d. *Context: When I get home from visiting my in-laws in Chile, I realize I left my sweater behind somewhere. It’s not at my husband’s parents’ place, so I don’t think I’ll find it again. I tell you:*

čkʷa xʷaʔaguxʷ šetʰ tekunukʷtoł.
 č=kʷa=xʷaʔag-əxʷ šə=tʰ=takinukʷt-ul

1SG.SBJ=CLD=LOSE-NCTR DET=1SG.POSS=sweater-PST

‘I lost my sweater.’ (vf | EP.2022/01/21)

When the past marker is attached to adjectives, the resulting interpretation is either that the adjective property has ceased to hold, as in (7c), or that the referent has ceased to exist (10B).

- (10) *Context: My husband and I have a multi-colored set of glasses. My favourite was the red one but it broke, and we threw it away. You’re putting glasses on the table before a meal and admire the remaining glasses in the set.*

A: hɛhɛw ʔajumišmot θ kʷoskʷasta.
 hihiw ʔaj-umiš-mut θ=kʷəs~kʷaʔsta
 really good-appearance-INT 2SG.POSS=PL~cup
 ‘Your cups are really beautiful.’

B: hɛl ʃɛ tətʰɛmɔl ʔə kʷɛhɛt
 hił ʃə=tətʰim-ul ʔə=kʷih-ít
 COP DET=red-PST CLF.PRT=increase-STAT
 ʔisxʷanoł ʔi kʷa yɪp.
 ʔəy-sxʷ-an-ul ʔiy kʷa=yəp̚
 good-CAUS-1SG.ERG-PST CONJ CLD=get.broken
 ‘The red one was my favourite but it broke.’
 (sf | EP.2023/06/29)

The use of the past marker on stative predicates (including nouns, adjectives, and stative verbs) in both clausal and nominal environments triggers an inference that the predicate bearing the past marker does not hold of its subject at the UT. For instance, the most natural interpretation of (7c) is that the dress is not white at the UT, while the referent of (9a) is understood to no longer be a house (and therefore no longer to exist). Following Thomas (2014) (who in turn takes the term from Altshuler and Schwarzschild 2012), I label this the *cessation inference*.

While it is beyond the scope of this paper to provide a full account of how this inference arises, I will sketch the analysis proposed in Thomas (2014). Briefly and informally, the idea is that a tensed proposition is interpreted in relation to contextually-relevant tensed alternatives: the assertion of the proposition is strengthened to mean that other contextually-relevant tensed alternatives do not hold (provided they are not entailed by the proposition). If a past tense proposition is asserted, the alternative present tense proposition is understood not to hold, so long as it is contextually relevant.⁵

The cessation inference does not arise when the context sets up a past topic time that does not include the present, as in (11). Here, the past topic time is the time of the speaker moving into the area. Note that (11) is a cleft and *hegus* ‘chief’ is the main predicate in the remnant clause.

⁵ See Altshuler and Schwarzschild (2012) for an alternative account of how the cessation inference arises.

- (11) *Context: Peter has been chief for a long time. I remember that he was chief when I moved into the area many years ago, and he still is today. I'm telling someone newer to the area.*

hɛl ʔot Peter ʔə hegusol šɛt⁰ ʔot q^wol
 hiɫ=ʔut Peter ʔə=higus-ul šə=t⁰=ʔut q^wəl
 COP=EXCL Peter CLF.PRT=chief-PST DET=1SG.POSS=at.first come
 tayqitol ʔə tɛʔɛ.
 tayq-iyt-ul ʔə=tiʔi
 MOVE-PRF-PST OBL=DEM
 ‘Peter was chief when I first moved here.’ (vf | EP.2023/07/07)

When the past suffix is used in nominal environments, however, cessation of the nominal property is entailed. The past suffix is infelicitous on *tičɛ* ‘teacher’ in (12a) and *laplɛt* ‘priest’ in (12b) because in each case, the nominal predicate still holds of the DP’s referent.

- (12) a. *Context: There’s a teacher that’s been at the school as long as we can remember, and he still hasn’t retired.*

hɛhɛw χ^woχ^wmot ʔəl nišəs šɛn
 hihiw ʃ^wuʃ^w-mut ʔəl=niš=as šan
 really long-INT COMP=be.here=3SBJV DEM
 tičɛh(#ol) ʔi qəʃi ʔot niš.
 tiča-h-(#ul) ʔiy qəʃi=ʔut niš
 teacher-EPEN-(#PST) CONJ still=EXCL be.here
 ‘That teacher has been here a long time, and he’s still (teaching) here.’ (sf | EP.2023/07/07)

- b. *Context: I see a news article about a parish priest in a small town where I used to live.*

nɛʔol tañ laplɛt(#ol) ʔək^w Duncan
 niʔ-ul tañ laplit-(#ul) ʔə=k^w=Duncan
 be.there-PST DEM priest-(#PST) OBL=DET=Duncan
 šɛt⁰ nɛʔol χ^woχ^wmotol. hɛhɛw
 šə=t⁰=niʔ-ul ʃ^wuʃ^w-mut-ul. hihiw
 DET=1SG.POSS=be.there-PST long.time-INT-PST really
 χ^woχ^wmotol ʔəl nɛʔəs.
 ʃ^wuʃ^w-mut-ul ʔəl=niʔ=as.
 long.time-PST COMP=be.there=3SBJV
 ‘That priest was in Duncan when I lived there long ago. He has been there a long time.’ (sf | EP.2023/08/31)

While this could be taken to indicate that the past marker in nominal environments is a distinct morpheme from the past marker in clausal environments, I do not pursue this approach. Tonhauser (2006, 2007) and Thomas (2012, 2014) observe a similar asymmetry for the past morpheme in Paraguayan Guaraní and Mbyá, but Thomas argues that this asymmetry arises due to different pragmatic factors in the interpretation of clauses and DPs, rather than to semantically distinct but homophonous past morphemes (one applying in nominalized clausal environments and one to nouns).

Briefly, Thomas analyzes the past marker on a noun as placing the RT for the nominal property in the past of the NP evaluation time. So, for instance, in (9b), the RT for *tičɛ* ‘teacher’ is placed in the past of the evaluation time for the NP *tičɛhoł* ‘former teacher’, which in this case is the same as the RT of the clause: the time of the seeing event.

Thomas proposes that the NP evaluation time is always relevant to the interpretation of the NP (i.e., it is always topical), and therefore the past marker on a noun always gives rise to the cessation inference: it is understood that the nominal property cannot be claimed to hold at the NP evaluation time.⁶ In (9b), this means that the referent of the DP *šɛ tičɛhoł* ‘a former teacher’ is understood not to be a teacher any longer at the NP evaluation time, the time of the seeing event.

Given the availability of a plausible pragmatic account, I believe a unified analysis of the past tense marker in nominal and clausal environments is preferable. At the very least, the presence of the same past marker applying across different environments with parallel interpretive differences in unrelated languages suggests that there should be a more general explanation than accidental homophony of nominal and clausal temporal markers.

If we adopt Thomas’s account, the different interpretations that arise when the past occurs in nominal environments can be understood in terms of the obligatory cessation inference. The following discussion follows Burton (1997) very closely, who also derives the various readings in terms of a cessation inference (though he does not use this term).

When the past occurs on an individual-level nominal predicate in a DP, the individual-level predicate is interpreted as ceasing to hold of the referent of the DP by the NP evaluation time. The resulting interpretation

⁶ Cable (2017) also shows that cessation inferences that arise for clausal uses of the optional past tense morpheme in Tlingit should be derived pragmatically, though they are often not cancellable.

is that the referent of the DP has ceased to exist by the time the NP is evaluated, as in (9a), since this is the only plausible way for the permanent property (being a house) to cease to hold of the referent (see also (7) and (10)).

When the past occurs on a stage-level noun like an occupation, the interpretation is that this temporary property no longer holds of the referent at the NP evaluation time. In this case, the meaning is compatible with the individual leaving the occupation, as for the retired teacher in (9b).

Finally, when the past occurs on a possessed noun, the RT for the possession relation and nominal property are placed in the past of the NP evaluation time. Following Burton (1997), the possession relation can be represented as a predicate R which is conjoined with the nominal predicate and has a possessum, possessor, and time argument: ... $N(x,t) \wedge R(x, \text{possessor}, t)$...⁷ It is the conjoined possession relation and nominal property that ceases to hold by the NP evaluation time. The cessation inference is met so long as one or both of the conjuncts cease to hold, resulting in both interpretations where the possession relation no longer holds and where the entity has ceased to exist, depending on context and plausibility.

If it is the possession relation that ceases to hold, the interpretation may be that the possession has been lost, as for the sweater in (9c), sold, or stolen. In contrast, since my grandparents will always stand in a grandparent relation to me, the past marker on *t^θ čičiyε?* ‘my grandmother’ and *t^θ k^wuk^wpa?* ‘my grandfather’ in (7b) results in the interpretation that these individuals are deceased.

6 The past marker *-oł* is not a tense

So far, the discussion of *-oł* leaves it plausible that it is a canonical past tense, placing the RT for the (verbal, nominal, or adjectival) predicate preceding a temporal anchor, and giving rise to cessation inferences in pragmatically determined contexts. However, there is one major difference between *-oł* and a morpheme that specifically occupies T: *-oł* does not have a fixed syntactic position and does not apply at a fixed point in the semantic derivation.

This is seen in complex nominal predicates (CNPs). CNPs consist

⁷ Burton (1997) has a separate possession time and nominal RT but I do without an extra possession RT here.

of a head noun preceded by one or more modifiers which together form the main predicate of the clause. The past marker can be found attaching both to an adjectival modifier and the main predicate, but the position of the past marker affects the interpretation.

When the adjectival property no longer holds, *-ot* attaches to the adjective. For instance, in (13a) and (14a), *-ot* attaches to the adjective *titolmot* ‘very small’. Because the lake and house still exist, *-ot* cannot felicitously attach to the head noun: (13b) and (14b) are infelicitous.⁸

(13) *Context: Daniel is pointing on a map to a little pond that used to be a big lake but was drained a while back for farmland.*

- a. **tihmotot** θayɛt taŋ sɣ^wox^wot ʔi
 tih-mut-**ut** θayal tin sɣ^wux^w-ut ʔiy
 big-INT-PST lake DEM long.time-PST CONJ
 gaʔqoθetəm.
 gəq-uθi[n]-t-əm
 open-mouth-CTR-PASS

‘This used to be a big lake a long time ago but they drained it.’

- b. #**tihmot** θayɛt**ot** taŋ sɣ^wox^wot ʔi
 tih-mut θayal-**ut** tin sɣ^wux^w-ut ʔiy
 big-INT lake-PST DEM long.time-PST CONJ
 gaʔqoθetəm.
 gəq-uθi[n]-t-əm
 open-mouth-CTR-PASS

‘This used to be a big lake a long time ago but they drained it.’

(sf | EP.2023/07/23)

(14) *Context: We’re looking at my neighbour’s house that used to be small but has had a lot of additions and renovations and is now quite big. I tell you:*

- a. **titolmotot** ʔaʔyeʔ taŋ ʔi paʂetəm.
 titul-mut-**ut** ʔa<ʔ>yɛʔ taŋ ʔiy paʂ-at-əm
 small-INT-PST house<DIM> DEM CONJ add.on-CTR-PASS

‘That used to be a small house, but they’ve added onto it.’

⁸ The fact that *-ot* is felicitous only on the adjective and not on the diminutive noun in (14) suggests that the contribution of the diminutive reduplication is not at-issue.

- b. #titolmot ʔaʔyeʔol tañ ʔi ʔašetəm.
 titul¹-mut ʔa<ʔ>yεʔ-**ul** tañ ʔiy ʔaš-at-əm
 small-INT house<DIM>-PST DEM CONJ add.ON-CTR-PASS
 ‘That used to be a small house, but they’ve added onto it.’
 (sf | EP.2023/07/16)

When the described entity no longer exists (and therefore both the nominal and adjectival properties no longer hold of it), the past marker can appear on either the adjective or the noun (15)–(17). My consultant sometimes preferred *-ol* on the adjective, but also accepted placement on the noun, unlike for (13)–(14); this preference is indicated with a question mark for (15b) and (16b).

(15) *Context: I point out an empty building in town to Daniel and Gloria:*

- a. ʔimot**ol** ʔeltənwtx^w tita sɣ^woɣ^wol.
 ʔəy-mut-**ul** ʔiltən-awtx^w təy^{ta} sɣ^wuɣ^w-**ul**
 good-INT-PST eat-building DEM long.time-PST
 ‘That used to be a good restaurant a long time ago.’
- b. ?ʔimot ʔeltənwtx^w**ol** tita sɣ^woɣ^wol.
 ʔəy-mut ʔiltən-awtx^w-**ul** təy^{ta} sɣ^wuɣ^w
 good-INT eat-building-PST DEM long.time-PST
 ‘That used to be a good restaurant a long time ago.’
 (sf | EP.2023/07/23)

(16) *Context: I’m showing you my yard and point out a stump.*

- a. hɛhew tihmot**ol** jεʔje tiʔta ʔi jεqatəm.
 hihiw tih-mut-**ul** jaʔja təy^{ta} ʔiy jaq-at-əm
 really big-INT-PST tree DEM CONJ fall-CTR-PASS
 ‘That used to be a big tree, but it’s been felled.’
- b. ?hɛhew tihmot jεʔjeh**ol** tiʔta ʔi jεqatəm.
 hihiw tih-mut jaʔja-h-**ul** təy^{ta} ʔiy jaq-at-əm
 really big-INT tree-EPEN-PST DEM CONJ fall-CTR-PASS
 ‘That used to be a big tree, but it’s been felled.’
 (sf | EP.2023/07/23)

(17) *Context: Peter was a good leader. He's since passed on.*

a. ʔimotɔɫ ʔəms heɡus Pita.
 ʔəy-mut-**uɫ** ʔəms=higus Pita
 good-INT 1PL.POSS=chief-PST Peter
 ‘Peter was a good former chief.’

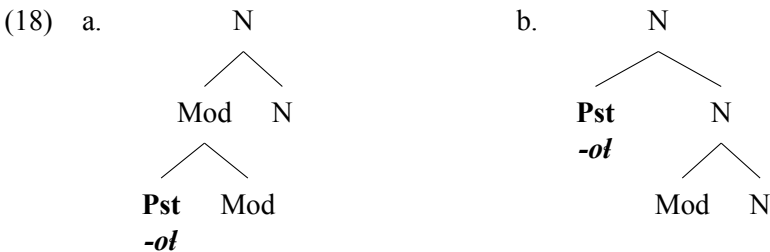
b. ʔimot ʔəms heɡusɔɫ Pita.
 ʔəy-mut ʔəms=higus-**uɫ** Pita
 good-INT 1PL.POSS=chief-PST Peter
 ‘Peter was a good former chief.’

(sf | EP.2023/07/19)

The final possibility where property of the head noun no longer holds of the subject, but the modifier property does, still needs to be investigated. The prediction is that the past marker will only be able to occur on the head noun in these cases.

This evidence from CNPs shows that the past marker can attach at different points syntactically, corresponding to differences in interpretation. To account for these facts, I propose that the past marker does not occupy T, and as such, does not constitute a true tense marker. Instead, following Matthewson (2005), I propose that the past marker is a temporal modifier.

When the modifier of a noun ceases to hold, as in (13)–(14), I propose that the past marker attaches directly to the modifier (18a). When both the nominal and modifier properties cease to hold, as in (15)–(17), I propose that the past marker attaches to the whole complex (18b).



Since modifiers typically precede the head of the phrase in ʔayʔajuθəɲ, I represent the past marker on a leftward branch. I assume that its suffixal specification causes it to attach to the head post-syntactically in the morphology (see Huijsmans 2022:99–109). When it is merged with a phrase, as in (19c), I suggest that it may either attach to the head noun or

to the closer preceding modifier; I leave a full account of how these two possible placements are derived to future work.

7 Semantic analysis

The common factor to the use of *-ot* across environments is that it restricts the RT for the verbal, nominal, or adjectival predicate to the past of a temporal anchor. While the temporal anchor in matrix clauses is typically the UT, the temporal anchor in nominal environments is the NP evaluation time.

I propose that the past marker contributes a presupposition that the RT t for a predicate P precedes the contextually provided temporal anchor t_c .⁹ In clausal contexts, the temporal argument will be saturated by the null non-future tense. In nominal contexts, I assume that the temporal argument is supplied by a contextually provided NP evaluation time.

$$(19) \quad \llbracket -ot \rrbracket^{c,g} = \lambda P.\lambda x.\lambda t : t < t_c.[P(x)(t)]$$

Crucially, since the past marker does not occupy T, its presence modifying NPs does not shed light on whether the NP evaluation time is syntactically represented. The past marker does not provide evidence for nominal tense.

This analysis predicts *-ot* to be able to combine with CNPs as a whole or their component parts. However, this analysis also raises questions, since the components of a CNP combine to take a single reference time supplied by T. Why then does it matter where *-ot* attaches? I sketch only a preliminary account here. A more complete analysis will require an account of how the component parts of the CNP combine semantically and is left for future research.

When the past attaches to the modifier in (13)–(14), it adds a presupposition that the RT for the modifier is in the past of the UT, triggering a cessation inference since the current states of the lake and house are salient and relevant. Since the RT for the complex predicate is ultimately saturated by the null non-future tense, the clause's RT ends up restricted to the past (consistent with the temporal adverb *sχ^woχ^wot* 'a long time ago' in (13)). *-ot* is infelicitous on *θay^let* 'lake' and *ʔaʔyeʔ* 'house (diminutive)'

⁹ The fact that different noun phrases within a clause could potentially have different temporal anchors means that having the temporal anchor provided as a parameter of interpretation is an oversimplification. However, providing a full account of the temporal interpretation of DPs is beyond the scope of this paper.

because triggering the presupposition in this position would signal that the lake property and house property required a past RT, contrary to fact.

For examples like (15)–(17), both the adjectival and nominal property have ceased to hold of the referent (since the restaurant no longer exists and chief has passed). For these cases, I have proposed that the past marker attaches to the whole CNP, meaning that it contributes a presupposition that the RT for the whole CNP is in the past. The cessation inference is then that the referent of the subject DP can no longer be described by the complex predicate.

8 Discussion

In this short paper, I have argued that the past marker in ʔayʔajuθəm is not a true tense, but rather a temporal modifier that can attach at different points within a clause. In terms of the debate regarding nominal tense in Salish languages, this paper supports the position taken by Matthewson (2005) where past markers in nominal contexts are modifiers rather than true past tenses or realizations of an interpretable tense feature (cf. Burton 1997; Wiltschko 2003). The past marker in ʔayʔajuθəm therefore does not provide evidence for a T projection among the functional projections of a noun phrase.

Though the past marker does not provide evidence for a tense projection in nominal environments, it does provide further evidence that the semantics of noun phrases involves reference to time. How the relevant temporal arguments are ultimately supplied is a matter for future research.

In closing, I would like to point to a welcome consequence of the current analysis. Besides accounting for why the past marker has variable placement, this proposal has the advantage of offering an explanation for why the past marker is not obligatory when the RT is past. Though the past marker contributes a presupposition that the RT precedes the UT, as a temporal modifier, it is not in competition with the null non-future tense. Therefore, even though it carries more presuppositional content than the null non-future tense, Maximize Presupposition does not apply (Heim 1991; Bochnak 2016) and the past marker is correctly predicted to be optional.

References

- Altshuler, D. and Schwarzschild, R. (2012). Moment of change, cessation implicatures and simultaneous readings. In Homer, V., Chemla, E., and Winterstein, G., editors, *Proceedings of Sinn und Bedeutung 17*, pages 45–62.
- Bochnak, M. (2016). Past time reference in a language with optional tense. *Linguistics and Philosophy*, 39:247–294.
- Burton, S. (1997). Past tense on nouns as death, destruction and loss. In Kusumoto, K., editor, *Proceedings of NELS 27*, pages 65–77, University of Massachusetts, Amherst. GLSA.
- Cable, S. (2017). The implicatures of optional past tense in Tlingit and the implications for ‘discontinuous past’. *Natural language and linguistic theory*, 35:635–681.
- Enç, M. (1981). *Tense without scope: An analysis of nouns as indexicals*. PhD thesis, University of Wisconsin.
- Enç, M. (1986). Towards a referential analysis of temporal expressions. *Linguistics and Philosophy*, 9:405–426.
- First Peoples’ Cultural Council (FPCC) (2022). Report on the status of B.C. First Nations languages. Retrieved from <https://fpcc.ca/wp-content/uploads/2023/02/FPCC-LanguageReport-23.02.14-FINAL.pdf>.
- Heim, I. (1991). Artikel und Definitheit. In von Stechow, A. and Wunderlich, D., editors, *Semantics: An international handbook of contemporary research*, pages 487–535. De Gruyter, Berlin.
- Huijismans, M. (2022). *Second-position clitics, from morphosyntax to semantics: The ʔayʔajuθəm (Comox-Sliammon) perspective*. PhD thesis, University of British Columbia, Vancouver.
- Huijismans, M. (2023). Lexical categories in ʔayʔajuθəm. In *Proceedings of the 58th International Conference on Salish and Neighbouring Languages*, Vancouver. UBCWPL.
- Huijismans, M. and Mellesmoen, G. (2021). An overview of control and non-control in ʔayʔajuθəm (Comox-Sliammon). In Reisinger, D., Green, H., Griffin, L., Huijismans, M., Mellesmoen, G., and Trotter, B., editors, *Proceedings of the 56th International Conference*

- on *Salish and Neighbouring Languages*, pages 84–110, Vancouver. UBCWPL.
- Klein, W. (1994). *Time in language*. Routledge, London.
- Matthewson, L. (2005). On the absence of tense on determiners. *Lingua*, 115:1697–1735.
- Matthewson, L. (2006). Temporal semantics in a superficially tenseless language. *Linguistics and Philosophy*, 29:673–713.
- Thomas, G. (2012). *Temporal implicatures*. PhD thesis, Massachusetts Institute of Technology.
- Thomas, G. (2014). Nominal tense and temporal implicatures: evidence from Mbyá. *Natural Language Semantics*, 22:357–412.
- Tonhauser, J. (2006). *The temporal semantics of noun phrases: evidence from Guaraní*. PhD thesis, Stanford University.
- Tonhauser, J. (2007). Nominal tense? the meaning of guaraní nominal temporal markers. *Language*, 83:831–869.
- Truth and Reconciliation Commission of Canada (TRC) (2015). Honouring the truth, reconciling for the future. Summary of the final report of the Truth and Reconciliation Commission of Canada.
- Underhill, R. and Cable, S. (2015). Hawaii trip. *Totem Field Storyboards*. Retrieved from <http://www.totemfieldstoryboards.org>.
- Watanabe, H. (2003). *A morphological description of Sliammon, Mainland Comox Salish, with a sketch of syntax*. ELPR Publications Series AZ-040. Osaka Gakuin University, Osaka.
- Wiltschko, M. (2003). On the interpretability of tense on D and its consequences for case theory. *Lingua*, 113:659–696.

The passive in Heritage Icelandic*

KRISTÍN M. JÓHANNSDÓTTIR
University of Akureyri

1 Introduction

Research has established that syntax tends to be rather well preserved in heritage languages, but that case marking is vulnerable and that an extensive use of the unmarked case is common (Benmamoun et al. 2013). This has been shown for instance for both Russian and Spanish heritage languages, as the inherent subject dative is replaced by the nominative (Benmamoun et al. 2013). This seems also to be the case in Heritage Icelandic, which has shown increased tendency to replace oblique subjects with nominative subjects (Arnbjörnsdóttir 2006; Óskarsdóttir & Bráinsson 2017; Björnsdóttir 2018).

So, what about constructions that require syntactical change and are morphologically complex, such as the passive voice? In many languages, the transformation from active to passive voice includes word order change. The functions of subject and object are reversed, making it more difficult to parse and produce (see e.g., Rodriguez et al. 2017). In Icelandic, the morphology is also complex as the auxiliary and the past participle behave differently based on the case, requiring two different strategies; dynamic features of number, person, and gender come into play with structural case but do not with lexical case. Because of this complexity, passives cross-linguistically are acquired rather late in first language, generally around the age of 6 to 7 (e.g., Marinis 2007; Kirby 2010). This is also the case in Icelandic as research shows that Icelandic pre-school children have difficulties comprehending the passive (Sigurjónsdóttir 2015). As Tsimpli (2014) has pointed out, late acquired phenomena are often more affected by reduced input, and the passive can therefore be difficult for heritage speakers, and research has shown that

* When it was time to write my thesis I couldn't choose between Hotze and Lisa. I knew they were both great and that they would both contribute enormously to my work, making it better and making me better. However, they had somewhat different styles and approaches, and I knew that I needed both to succeed. I had taken classes from Hotze, and I had been his teaching assistant, so I knew he would be supportive and encouraging, but I also knew he would push me, not let me get away with slacking off and scold me when needed — and I knew I needed that. Fortunately, they agreed to co-supervise me and for that I'm forever grateful. Thanks for all you did for me, Hotze, and you Lisa both.

young heritage speakers of Spanish simply avoid the passive (Silva-Corvalán 2014). So, what is the status of the passive in Heritage Icelandic, a language spoken by the descendants of Icelandic immigrants to North America, bilinguals who live in an English-dominant language environment?

The goal of this paper is to explore the passive voice in Heritage Icelandic as it pertains to the syntactic structure and case assignment, to see whether the passive is stable or vulnerable in the language. For this purpose, 29 speakers of Heritage Icelandic participated in a task-specific examination, where they had to choose their preferred version of a passive sentence. Results show that the syntactic structure of the passive is still rather strong, as suspected, but that there are definite signs of attrition in case marking.

2 The passive in Icelandic

The passive in Icelandic is formed with the auxiliary verb *vera* ‘be’ plus a past participle of the main verb:¹

- (1) Jón var kysstur.
 John was.PAST kissed.PAST PART
 ‘John was kissed.’

The subject of the passive sentence corresponds to the object of a similar active voice sentence. It is base generated in object position and moved to a specifier position with an NP-movement (e.g., Þráinsson 2005). Case assignment then depends on the case. If the original object is not assigned a lexical case, it is assigned structural case in its object position, which in Icelandic is the accusative. Then, when it is passivized into subject position, it is assigned the nominative.

- (2) a. Jón barði **Guðmund**.
 Jón beat Guðmundur.ACC
 ‘John beat Guðmundur.’

¹ The following abbreviations are used in this paper: ACC = accusative, DAT = dative, DET = determiner, GEN = genitive, NOM = nominative, NT = neuter gender, P = person, PAST = past tense, PAST PART = past participle, PL = plural, SG = singular.

- b. **Guðmundur** var barinn.
 Guðmundur.NOM was beaten
 ‘Guðmundur was beaten.’

If the original object is assigned a lexical case (dative or genitive), it does not undergo a change in case when passivized, resulting in an oblique subject.

- (3) a. Maðurinn ýtti **Guðmundi**.
 man.DET pushed Guðmundur.DAT
 ‘The man pushed Guðmundur.’
- b. **Guðmundi** var ýtt.
 Guðmundur.DAT was pushed
 ‘Guðmundur was pushed.’
- (4) a. Hann saknaði **Guðmundar**.
 he missed Guðmundur.GEN
 ‘He missed Guðmundur.’
- b. **Guðmundar** var saknað.
 Guðmundur.GEN was missed
 ‘Guðmundur was missed.’

In addition to this different behaviour in case assignment, there is also difference in agreement. With structural case the subject and the verb agree in case whereas with lexical case they do not; the auxiliary is always in third-person singular and the past participle in the neuter.

- (5) a. Konunni var ýtt.
 woman.DET.3P.SG was.3P.SG pushed.NT
 ‘The woman was pushed.’
- b. Mér var ýtt.
 I.1P.SG was.3P.SG pushed.NT
 ‘I was pushed.’
- c. Ykkur var ýtt.
 you.2P.PL was.3P.SG pushed.NT
 ‘You were pushed.’

In Icelandic, the agent is usually not present in a passive sentence although it can be introduced in a prepositional phrase:

- (6) Konunninni var ýtt af mannum.
 woman.DET was pushed by man.DET
 ‘The woman was pushed by the man.’

One of the most noticeable changes in Icelandic syntax in recent years is the so called “New Passive”, first mentioned in print in 1979 (Jónsson 1979). In the “New Passive”, the object doesn’t raise to the subject position but stays in object position and keeps its case, whether it is structural or lexical. Instead, the expletive *það* ‘it’ is inserted into the subject position:

- (7) a. Jón barði mig.
 John beat me.ACC
 ‘John beat me.’ [Active]
- b. Ég var barinn.
 I.NOM was beaten
 ‘I was beaten.’ [Passive]
- c. Það var barið mig.
 it was beaten me.ACC
 ‘It was beaten me’ = ‘I was beaten.’ [“New Passive”]

This “New Passive” construction is extremely common among young people but hardly used by anyone over the age of 30 (see e.g., Sigurjónsdóttir & Maling 2001; Sigurjónsdóttir 2018). The reasons for the change are not clear but this new construction is obviously much simpler than the original passive as it does not require NP-movement of the object, nor a change in case; the object stays in-situ, in its original case.

When comparing the Icelandic passive to the English passive, we see the same NP-movement of the object from object position to the subject position, and when the object/subject is the first-person pronoun, we even see a change in structural case (8), something that is not visible in any other instance (9):

- (8) a. John beat me.
 b. I was beaten by John.

- (9) a. John beat **the woman**.
b. **The woman** was beaten by John.

As English has all but lost its case marking, speakers of English do generally not have to think about a change of case in their passives; this is something that might affect the passives in Heritage Icelandic. We might therefore expect the syntactic structure of the passive to be rather stable in Heritage Icelandic but the case assignment to be affected.

3 The current research

Heritage Icelandic is a language spoken by the descendants of Icelandic immigrants to North America from approximately 1870 till the First World War. The speakers today are generally third or fourth generation speakers and they are getting old; many of them haven't used the language on regular basis for decades. Of those speakers that can still be considered fluent speakers of the language, most live in Manitoba, Canada, particularly in the area called New Iceland. There are also still some speakers in Saskatchewan and North Dakota, but fluent speakers in other provinces or states are usually people that grew up on the prairies and moved away during their adult years.

Data used in this study were elicited in 2014 from 29 speakers, 13 men and 16 women. Of these 29, 14 were from Manitoba, 4 from Saskatchewan, and 8 from North Dakota. Average age was 75.18, ranging from 35 to 97. All but the youngest speaker had spoken Icelandic from birth, but it varied at what time English took over as the dominant language. In general, the people from Manitoba only spoke Icelandic until they went to school around the age of six, but the speakers from Saskatchewan and North Dakota were more likely to have been exposed to English earlier, even from birth. Considering that children do not acquire the passive until around 6 to 7, as previously stated, this means that English had already taken on a leading role for most speakers by the time they had fully acquired the passive in their heritage language.

The speakers took a judgement test where they were asked to choose acceptable sentences from a list of various passive sentences. First, they would see a context sentence and then three or four options were given

for the test sentence.² These sentences would vary in case and in syntactical structure.

(10) *Það kom upp leiðinlegt atvik.*

'There was an unpleasant incident.'

- a. **Stráknum** var hrint. 'The **boy (DAT)** was pushed.'
[Regular passive]
- b. **Strákurinn** var hrintur. 'The **boy (NOM)** was pushed.'
[Passive with NOM]
- c. Það var hrint **stráknum**. 'It was pushed the **boy (DAT)**.'
[“New Passive”]
- d. Það var hrint **strákurinn**. 'It was pushed the **boy (NOM)**.'
[“New Passive” with NOM]

There were two different tests used, each including six passive sentences, three that required a nominative subject and three that required a dative subject. Genitive was not tested as very few verbs that assign genitive can be passivized and they are very rare in the language. They may therefore not be in the vocabulary of a heritage speaker.

In addition to differences regarding case, the test sentences differed in syntactical complexity such that two sentences had direct word order, two included an interrogative pronoun in subject position, requiring a V2-inversion, and two sentences included an AdvP or PP in subject position, which also requires a V2-inversion. Even though the syntactic structures of the passives in Icelandic and English are rather similar, English does not have V2-inversion, so the more complex of these sentences might cause the speakers some problems.

Based on what we now know, three predications are made for Heritage Icelandic:

Prediction 1: The passive construction is rather well preserved but there will be signs of confusion.

Prediction 2: The “New Passive” is not common.

² Three sentences were given for structural case as there were no sentences with an accusative subject. However, four sentences were given as an option when the original object was in the dative case.

Prediction 3: There will be a tendency for the nominative to replace oblique subjects.

4 Results

4.1 The syntax

As the syntax of the English passive doesn't differ much from that of Icelandic, and as syntax is generally not much affected in heritage languages, we didn't expect the syntax of the passive in Heritage Icelandic to show much sign of weakening. In fact, the correct syntactical structure, including an NP-movement, was chosen 76% of the time whereas the syntactical structure of the "New Passive", with the object in-situ, was chosen 24% of the time. Furthermore, eleven speakers, or 38% of all speakers, always picked the correct sentence structure, indicating no syntactical confusion with the passive for them.

However, 18 speakers (62%) did pick the in-situ version at least once, indicating that their passives may be at least somewhat shaky, and Rodriguez et al. (2017) showed that the passive can indeed cause heritage speakers some confusion, particularly in production. When the data from these 18 speakers are examined, we see not only inter-speaker variation but also intra-speaker variation as some of the speakers only picked the in-situ version once or twice whereas others seemed to prefer that version. Six speakers chose the in-situ version more often than the NP-movement version, although all but one chose the regular passive construction at least once. That speaker correctly chose the regular passive in the pre-test sentence but never in the actual test.

Interestingly enough, when we look at the syntactical structures of these passive sentences, one might have speculated that the heritage speakers would do better with simple constructions than those that require a V2-movement, particularly since there are signs of weakening of the V2-system in Heritage Icelandic (Arnbjörnsdóttir et al. 2018). However, the sentences that started with an interrogative pronoun, requiring a V2-inversion, had the highest accuracy rate, 86.4%, whereas the simple construction and the construction with an AdvP or PP in the initial position, also requiring a V2-inversion, had an accuracy rate just over 74%. There are therefore no signs of the simpler constructions faring better when it comes to the passive.

For some speakers the passive seems still strong.³ However, as 62% of the speakers show some signs of affected passive, and some of them seem to choose the in-situ version over the one that includes an NP-movement, there might be reasons to speculate that Prediction 1 might be underestimating the situation.

P1: The passive construction is rather well preserved but there will be signs of confusion.

The correct construction is picked 75% of the time, but 62% of the speakers nevertheless pick the incorrect construction at least once.
= P1 borne out?

Now, one might believe that 24% is a rather high number for problems with the syntax, particularly when the two languages are so similar, and that it might indicate that the “New Passive” also exists in Heritage Icelandic, but before we jump to conclusions, we should look at the in-situ sentences in more detail.

4.2 The “New Passive”

As discussed in Section 2, the so-called “New Passive” has become increasingly common in the language of younger Icelanders, and the fact that 62% of the speakers pick the in-situ-version at least once might indicate that a similar change is taking place in Heritage Icelandic, even though there has been no new immigration to talk of for over a hundred years. However, when we look at our heritage data, we see one major difference between Icelandic and Heritage Icelandic — case. The thing to remember here is that in the “New Passive”, the in-situ object includes no case change, meaning that a dative object stays dative, and an accusative object stays accusative, as seen in (7c). However, when we look at the in-situ sentences in Heritage Icelandic, we see quite a different pattern. In 73% of the cases, the speakers choose the nominative object and only in 27% of the instances the “New Passive” version with an accusative or dative object. This means that the “New Passive” construction is only chosen in 7% of the instances, by only nine speakers. What we have here can therefore hardly be the rise of a “New Passive”,

³ We should nevertheless keep in mind that Rodriguez et al.’s (2017) study indicated that recognizing the correct passive is easier than producing it correctly and as this study only required the speakers to pick out the correct passive, we might see stronger outcomes than if we asked the speakers to produce the passive.

like that in Icelandic, but instead we seem to have indications of an expanded use of the nominative case. Prediction 3 is therefore borne out:

P2: *The “New Passive” is not common.*

Only 7% of the sentences chosen are of that construction.

4.3 The case system

Case is much more vulnerable in heritage languages than syntax and, in the passive, the Icelandic case system is much more complex than that of English. One would therefore expect the passive in Heritage Icelandic to show more attrition regarding case.

In the sentences where a dative subject is expected, 68% of the speakers correctly pick the sentence with the dative subject and the nominative is only picked 32% of the time, indicating a rather strong position of the dative in passive sentences. However, as 16 speakers (55%) chose a nominative subject over a dative one at least once, it also tells us that just like with the syntax, there is both inter-speaker variation and intra-speaker variation when it comes to case.⁴

As we see strong signs of intra-speaker variation, it is interesting to see whether it is completely random when the dative is kept or whether there is any pattern to it; that is, whether the structure of the sentence might affect the choice of case. The data show that the dative keeps its position best in a direct word order, 76% accuracy rate, but least with an interrogative pronoun, 62% accuracy rate. The sentence structure with an AdvP or PP falls in between with 67% accuracy rate. The fact that V2-inversion didn't seem to affect the accuracy rate of the NP-raising might make it less likely that the complexity of the syntax is affecting case assignment here, and a much bigger data pool would be required for any such claim. However, Icelandic is a V2-language and V2 is more vulnerable in topicalization structures in Heritage Icelandic (Arnbjörnsdóttir et al. 2018), so we cannot rule it out that syntax is affecting the case assignment in some way.

The fact that 55% of the speakers chose a nominative subject over the dative subject when the syntax was otherwise correct might indicate a more general tendency for the dominance of the nominative which would be in line with English, a language that has all but lost its case marking. As we generally don't get any accusatives in the subject position of the passive in Icelandic, and the test sentences didn't account for any such

⁴ Only one speaker always picked the nominative subject.

cases,⁵ we can only compare sentences where the speakers chose the in-situ version. Here the results show that the nominative replaces the accusative in object position 77% of the time and the dative 64% of the time, meaning that the nominative replaces the accusative more often than the dative. As the accusative is seen as the unmarked case of the Icelandic object, we would have expected this to be the other way around. However, Arnbjörnsdóttir (2006), who reported a confusion in case marking in Heritage Icelandic, pointed out that there didn't seem to be any signs of it being regular or consistent and there was even a tendency of using the dative where there should be an accusative. Björnsdóttir (2018:355) reported a similar tendency of the dative replacing accusative and genitive objects. On the other hand, Dehé and Kupisch (2021) saw increase in the use of nominative and accusative case at the expense of the dative. These contradicting results indicate that the case system in Heritage Icelandic is quite vulnerable and that there is some confusion as to which case to use. Therefore, it seems that Prediction 3 is borne out:

P3: There will be a tendency for the nominative to replace oblique subjects.

55% of the speakers pick a nominative subject instead of dative subject at least once, indicating a clear tendency for the nominative to replace an oblique subject.

This is in line with other studies that show that case assignment is vulnerable in heritage languages and that the unmarked subject case has the tendency to replace the dative in subjects (Benmamoun et al. 2013). What we have here is perhaps even increased confusion with more complex syntax.

5 Conclusion

The syntax of the passive is rather strong in Heritage Icelandic, but it nevertheless shows some signs of vulnerability as to whether the object rises to subject position or not. There are also indications of changes in the case marking where the nominative seems to be overextending, not only as the subject case but in some instances also as an object case. This is in line with previous research that shows that the syntax of heritage

⁵ It would have been interesting to see if the speakers would ever have picked a sentence with an accusative subject, indicating that they treated the structural case in the same way as the lexical case, but such sentences were not included in the test.

languages is rather strong but that the case system is vulnerable (Benmamoun et al. 2013). When a heritage speaker is faced with a rather complex linguistic process, such as the passive, it is natural that they may show some vulnerability, not only in their production but also in comprehension. This may not necessarily mean a reduction in the case system but possibly new semantic domains. Even though these speakers may not perform exactly like the speakers of the base-language, Heritage Icelandic is a completely grammatical system which shows some signs of reanalysis of the structural system.

References

- Arnbjörnsdóttir, Birna. 2006. *North American Icelandic. The life of a language*. Winnipeg, MB: University of Manitoba Press.
- Arnbjörnsdóttir, Birna, Höskuldur Thráinsson, & Iris Edda Noweinstein. 2018. V2 and V3 orders in North-American Icelandic. *Journal of Language Contact* 11:379–412.
- Benmamoun, Elabbas, Silvina Montrul, & Maria Polinsky. 2013. Heritage languages and their speakers: Opportunities and challenges for linguistics. *Theoretical Linguistics* 39(3–4):129–181.
- Björnsdóttir, Sigríður Mjöll. 2018. “Hún er svo montin af að vera íslenskt.” Málbreytingar í sendibréfum Vestur-Íslendinga. *Sigurtunga. Vesturíslenskt mál og menning*. Reykjavík, Iceland: Háskólaútgáfan, 341–360.
- Dehé, Nicole & Tanja Kupisch. 2021. Prepositional phrases and case in North American (heritage) Icelandic. *The Nordic Journal of Linguistics* 45(3):254–280.
- Jónsson. Gísli. 1979. Íslenskt mál. 1. þáttur. *Morgunblaðið*, 20. May, p. 14.
- Kirby, Susannah. 2010. Passives in first language acquisition: What causes the delays? *University of Pennsylvania Working Papers in Linguistics* 16(1):109–117.
- Marinis, Theodoros. 2007. On-line processing of passives in L1 and L2 children. *Proceedings of the 2nd Conference on Generative Approaches to Language Acquisition North America (GALANA)*, 265–276.

- Óskarsdóttir, Salbjörg & Höskuldur Þráinsson. 2017. Hann þótti gott í stauþinu. Um breytingar á aukafrallsfrumlögum í vesturíslensku. In Höskuldur Þráinsson, Ásgrímur Angantýsson, & Einar Freyr Sigurðsson (eds.), *Tilbrigði í íslenskri setningagerð. III Sérathuganir*. Reykjavík, Iceland: Málvísindastofnun Háskóla Íslands, 101–124.
- Rodriguez, Estrella, Kristina Bustamante, Carla Wood, & Gretchen Sunderman. 2017. A comparison of the grammatical production of child heritage speakers of Spanish across language and grade: Kindergarten and grade 1. *Languages* 2(27).
- Silva-Corvalán, Carmen. 2014. *Bilingual language acquisition: Spanish and English in the first six years*. Cambridge Approaches to Language Contact. Cambridge, UK: Cambridge University Press.
- Sigurjónsdóttir, Sigríður. 2015. Acquisition of the new impersonal construction in Icelandic. In Elizabeth Grillo & Kyle Jepson (eds.), *Proceedings of the 39th Annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press, 385–397.
- Sigurjónsdóttir, Sigríður. 2018. Nýja þolmyndin nú og þá. Samanburður á milli tveggja kannana. In Höskuldur Þráinsson, Ásgrímur Angantýsson, & Einar Freyr Sigurðsson (eds.), *Tilbrigði í íslenskri setningagerð. III Sérathuganir*. Reykjavík, Iceland: Málvísindastofnun Háskóla Íslands, 241–272.
- Sigurjónsdóttir, Sigríður & Joan Maling. 2001. Ný setningafræðileg breyting? Um hina svokölluðu “nýju þolmynd” í íslensku. *Málfregni*, 20. URL: <http://hdl.handle.net/10802/982>
- Tsimpli, Ianthi. 2014. Early, late or very late? Timing acquisition and bilingualism. *Linguistic Approaches to Bilingualism* 4(3).
- Þráinsson, Höskuldur. 2005. *Setningar. Handbók um setningafræði*. Reykjavík, Iceland: Almenna bókafélagið.

Where to Interpret What*

KYLE JOHNSON

University of Massachusetts at Amherst

Hotze is known now for his work on tense, and on the vagaries of pronouns, but I got to know him when he was working on questions. His work on the semantics of *wh*-questions — a lot of it with Sigrid Beck — heavily influenced my thinking about the syntax of movement. I blame him for my resulting, years-long, obsession with multidominance. His work was the first step in a long line of interesting work on the syntax and semantics of *wh*-questions that continues today. The immediate predecessor to this work was Hotze's equally important dissertation: one of the first attempts to explain an island condition entirely from its semantics. It remains an important role model for the contemporary work on the semantics of islands, and opened my eyes to the wider possibilities of finding the source of islands. Thank you Hotze for starting me on a journey that has dominated my research life. But the reason I'm contributing to your volume is even more personal: it's because the other thing I learned when I got to know you is how much I like you.

In this note, I'll sketch a few facts about *wh*-movement that expand on the view in Beck and Rullman (1998) and Rullmann and Beck (1998) that *wh*-phrases are interpreted in their underlying position, no matter where they show up in the surface representation. In addition to the semantic reasons for this conclusion, there are straightforward facts about anaphora that animate this view. A famous kind of example of this is (1).

(1) Which articles about herself should no woman respond to?

This sentence has an interpretation in which *herself* is understood as a variable bound by *no woman*. Under normal circumstances, a reflexive pronoun in English, like *herself*, can only be bound by arguments that c-command them, are local, and sit in Argument positions. There is no successful definition of local and Argument position that I am aware of, but c-command is serviceably defined by (2).

* My thanks to Satoshi Tomioka for help on this paper.

(2) α c-commands β iff β (reflexively) dominates α 's sister.

These three requirements on binding a reflexive are illustrated by (3)-(5).

- (3) a. No woman₁ should respond to these articles about herself₁.
b. *No woman₁'s son should respond to these articles about herself₁.
- (4) a. No woman₁ should should ask herself₁ about the articles that respond to me.
b. *No woman₁ should ask me about the articles that respond to herself₁.
- (5) a. Each day₁ begins with its₁ best meal.
b. *Each day₁ I eat its₁ best meal at dawn.

(3b) violates the c-command requirement, as the sister to *no woman* is *son*, and *herself* is not included in *son*. (4) violates the locality condition—which requires very roughly that the reflexive be in all the CPs dominating its binder. In (4b) the relative clause is a CP that contains *herself* but not *no woman*. And finally, (5b) violates the requirement that the binder sit in an Argument position. Unlike (5a), *each day* does not sit in an Argument position in (5b), and for this reason differs from (5a) in being able to bind *its*.

What (1) shows, then, is that the position from which a wh-phrase moves can be used to calculate whether a reflexive it contains meets its requirements for being bound. Engdahl (1980) showed that functional readings for questions have a similar dependency on terms that c-command the position from which the wh-phrase moves. To see this, consider the pair of sentences in (6).

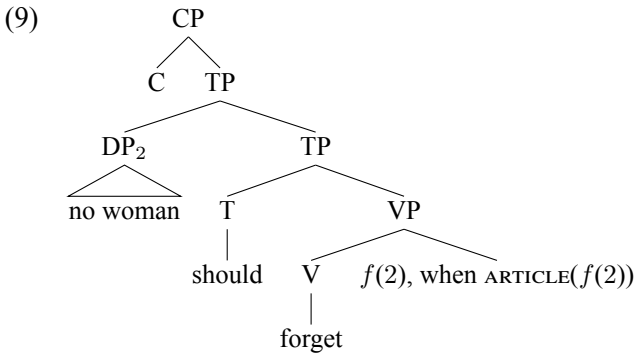
- (6) a. Which article should no woman forget?
b. Which article should the advisor no woman worked with forget?

The syntax and semantics for (6a) should allow (7) to be a short version of the answer in (8).

(7) her first article

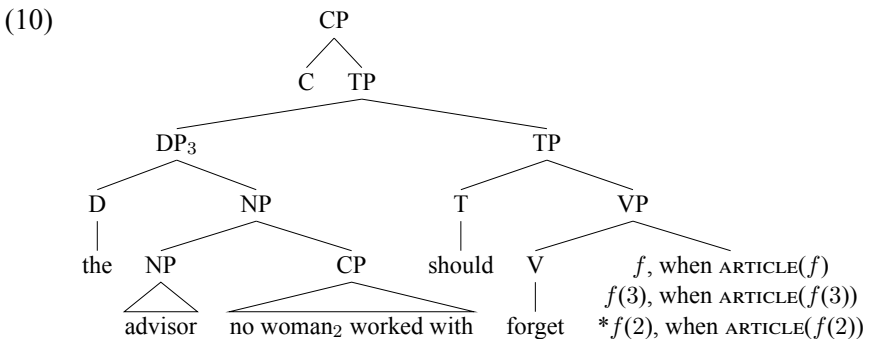
(8) No woman₂ should forget her₂ first article.

But the syntax and semantics for (6b) should disallow the answer (7) to express a parallel relation between *her* and *no woman*. Engdahl suggested that questions involve quantification over a Skolemized choice function. The variable within that choice function must be in the scope of its binder to deliver an interpretation where the function varies with the binder. Schematically, we need something like (9), for (6a). (*f* is semantic type $\langle e, e \rangle$.)



Let questions be sets of propositions that differ just in the value that *f* takes. In (9), *f*'s value depends on the value 2 has, and (7) is the answer (8) because it provides the value “ $\lambda x x$'s first article” for *f*.

By contrast, (6b) cannot involve an *f(x)* whose variable is bound by *no woman*. If it has a variable, that variable must be bound by the *advisor* DP. (Engdahl argues that the functions can have any number of variables. It can also be the constant, non-Skolemized, choice function *f*.)



For this reason, (7) can only be the shortened version of one of the answers in (11).

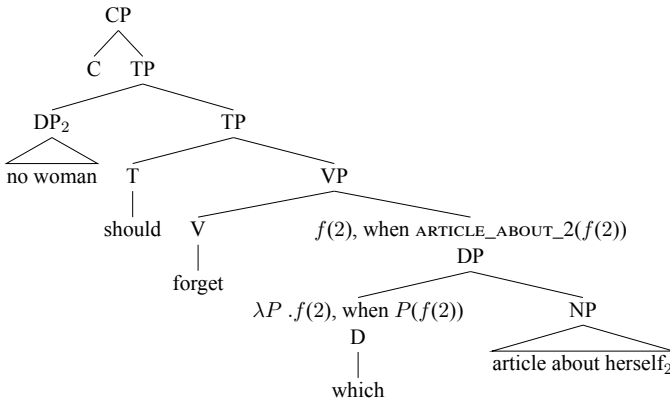
(7) her first article.

- (11) a. The advisor₃ should forget her₃ first article.
 b. The advisor should forget her first article. (*her* contextually given)

We can capture these two phenomena if the denotation for the material in the pre-moved position contains the syntactic material in the moved position. Let's start, then, with the assumption that the syntactic representation which is semantically interpreted is (12).

- (12) a. $[[\text{which}]] = \lambda P.f^*, \text{ when } P(f^*) (f^* \in \{f, f(x), f(x, y), \dots\})$

b.



This question seeks information about a function, f , which depends on the value given to 2 and selects an entity which is an article about 2. An appropriate answer to such a question is *the nicest* which will pick out, for each woman, the nicest article about that woman.

On this view, *wh*-determiners can have a hidden pronoun in them which is capable of being bound. Interestingly, this hidden pronoun is also subject to the requirement that the binder be in an Argument position. This is demonstrated by the contrast in (13).

- (13) a. Which meal should each day begin with?
 b. Which meal did you tell me each day to eat?

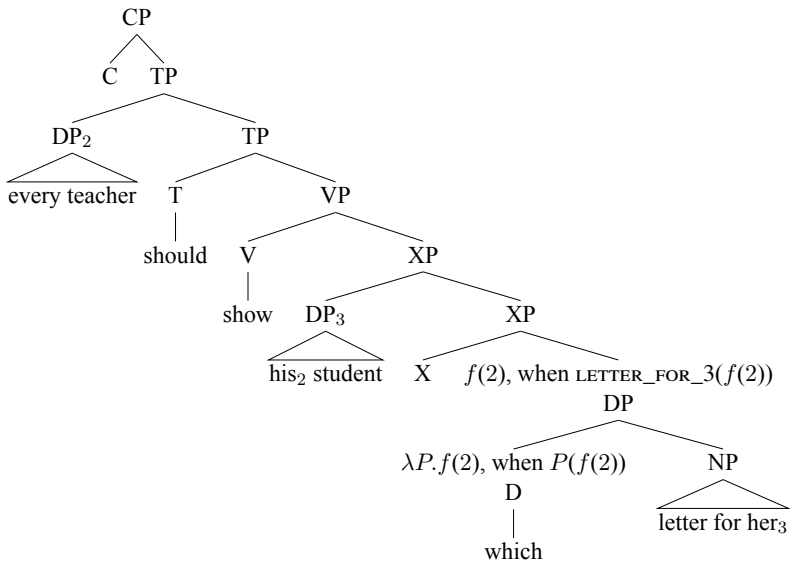
The answer *its best* is inappropriate for (13b) unless the context provides a value for *it*. Only as an answer to (13a) can it name a function whose variable, i.e. *it*, is bound by *each day*. Only in (13a) is *each day* in an

Argument position. The difference in (13) corresponds to the fact that only in (14) can *it* be bound by *each day*.

- (14) a. Each day₂ should begin with its₂ best meal.
- b. *You told me each day₂ to eat its₂ best meal.

This analysis allows the variable that comes with the choice function associated with *which* to be different than any variable that is within the NP, like *herself* in (12). This correctly predicts that questions like (15) can receive an interpretation in which these variables have different binders.¹

- (15) Which letter for her should every teacher show his student?



This question seeks the identity of the *f* which, for each teacher-student pair, picks the thing for the teacher that is a letter for the student. This allows *his best* to be a short form of the answer in (16).

- (16) Every teacher₂ should show his₂ student₃ his₂ best letter for her₃.

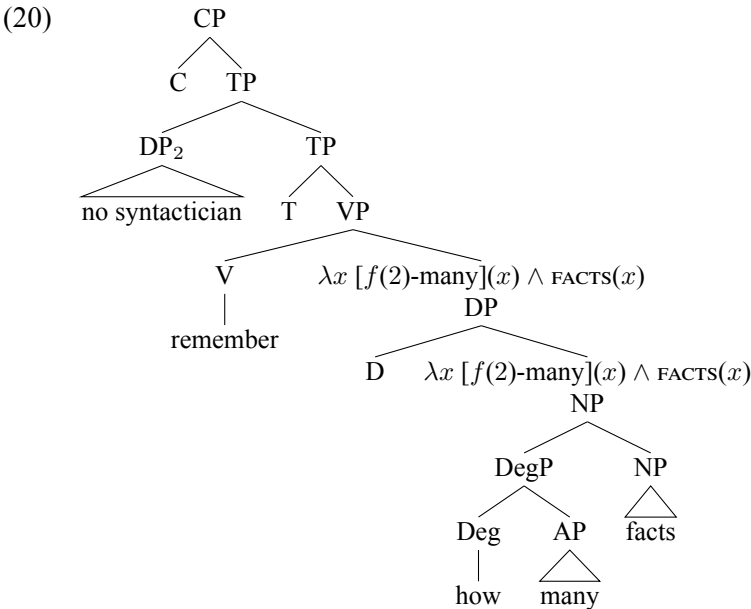
This account extends to questions involving degrees. Degree questions can also seek the identity of a Skolemized function. The expression in (18a) can express the answer to (17) that is found in (18b).

¹ My thanks to Satoshi Tomioka for help with this example.

- (17) How many facts should each syntactician remember?
- (18) a. no fewer than are needed to solve the problem at hand
 b. Each syntactician should remember no fewer facts than are needed to solve the problem at hand.

The dialogue in (17)-(18) allows there to be different solutions to the problem at hand. Those solutions can depend on the facts each linguist has in their memory. Christopher Hammerly, for instance, might need only to remember two facts he knows about Ojibwe to solve the problem at hand, whereas I, ignorant of those facts, can do no better than to remember four facts about Norwegian in forming my solution. The answer in (18) picks out numbers of facts that depend on the syntactician — two for Dr. Hammerly, four for me, etc. This can be derived if (17) has a representation something like (20).

(19) $\llbracket \text{how} \rrbracket = \lambda P \lambda Q \lambda x [f(y)-P](x) \wedge Q(x)$



The degree word *how* introduces an f whose values (=degrees) depend on the value that 2 gets. Supplying the answer in (18) will cause the object of *remember* to be two-many facts for Hammerly, and four-many facts for me. Just like *which*, then, *how* contains a Skolemized choice function.

As expected, both the c-command and Argument position conditions hold of *how* as well. (18a) doesn't provide an answer that maps syntacticians to numbers of facts if it is the answer to the question in (21).

- (21) How many facts should the linguist no syntactician likes remember?

That's because the syntax for (21) doesn't put *how* in the scope of *no syntactician*, and as a consequence *no syntactician* cannot bind a variable that *how* comes with.

Similarly, the contrast in (22) indicates that being in an Argument position is a prerequisite for a term to be a binder for the variable that comes with *how*.

- (22) a. How many hours of sunlight does every day contain?
 b. How many hours of sunlight did you tell me every day to get?

The expression *less than the previous day* can pick out times that depend on the value *every day* gets only in (22a); and only in (22a) is *every day* in an Argument position.

And finally, just as in *which* questions, the variable that comes with the choice function in *how* can have a different binder than a pronominal variable elsewhere in the wh-phrase. This is demonstrated by the question/answer pair in (23).

- (23) a. How many stories about himself₂ should no advisor₃ tell her₃ student?
 b. more than he needs to hear

The answer in (23b) names a function which provides degree-student pairs that vary by advisor. This happens because *how* contains a variable bound by *no advisor* — this delivers the degree part of the pair, one for each advisor — and *himself* is bound by *her student* — this delivers the student part of the pair, again, one for each advisor.

None of this would be true if Hotze (and Sigrid) weren't right all those years ago about where wh-phrases are semantically interpreted. What the conditions on binding of reflexives in (1) show is that this is because the syntax must, despite where the wh-phrase may be pronounced, put that wh-phrase in its interpreted position.

References

- Beck, Sigrid, and Hotze Rullmann. 1998. Reconstruction and the interpretation of *Which* phrases. In *Proceedings of the 1997 tuebingen workshop*, ed. Graham Katz, Shin-Sook Kim, and Heike Winhart, number 340 in *Arbeitspapiere des Sonderforschungsbereichs*, 223–256. Universities of Tuebingen and Stuttgart.
- Engdahl, Elisabet. 1980. The syntax and semantics of questions in swedish. Doctoral Dissertation, University of Massachusetts at Amherst.
- Rullmann, Hotze, and Sigrid Beck. 1998. Presupposition projection and the interpretation of *Which*-questions. In *Proceedings of Semantics and Linguistic Theory 8*, ed. Devon Strolovitch and Aaron Lawson, 215–232. Ithaca, NY: Cornell Linguistics Club.

Building statives in Nsyilxcn*

JOHN LYON

University of British Columbia – Okanagan

1 Introduction

Nsyilxcn (a.k.a. Colville-Okanagan, ISO 639-3: oka) is a Southern Interior Salish language spoken in south-central British Columbia, and the northern interior of Washington State. There now remain approximately 81 fluent first-language speakers in Canada (FPCC 2022), though there are intensive revitalization efforts underway on both sides of the border.

This paper develops a semantics for derived, stative predicates in Nsyilxcn. These predicates are formed by attaching a stativizer (*a*)*c*- to a change-of-state (CoS) root, as in (1).¹

- (1) a. **c-ǰay'** mnímtət i? stəłtált-(t)ət i? kł
STAT-get.written 1PL.INDP DET truth-1PL.POSS DET to
scəcmálaʔ-tət.
children-1PL.POSS
'Our family declaration is written.'
(Delphine Derickson Armstrong, VF)
- b. lut kn ta kł-kəwáp alí?
NEG 1SG.SUBJ NEG.FAC have-horse because
c-naǰ' in-kəwáp.
STAT-get.stolen 1SG.POSS-horse
'I don't have a horse because my horse is stolen.'
(Delphine Derickson Armstrong, VF)

* Thank you, Hotze, for opening my mind to the logic of language as a young Ph.D. student, and for your tireless support of semantic work with indigenous languages. Many thanks go to Delphine Derickson Armstrong of Westbank reserve, without whom none of this work would have been possible. Thanks also to *twi*-Lottie Lindley and Sarah McLeod of Quilchena. *talí? kʷu kʷukstp! Limləmt* to my research assistants Hailey Causton, Tish Elkink, and Ashley Gregoire. This work has been funded through SSHRC IDG 430-2022-00827, and supported by the En'owkin Centre. A longer, more-detailed version of this paper is published in the 58th Annual ICSNL Proceedings, available at UBCWPL's online Kinkade Archive. Contact: john.lyon@ubc.ca

¹ Abbreviations used in this paper are as follows: c2 – final (inchoative) reduplication; DET – determiner; DUB – dubitative; FAC – factual; INCH – inchoative; INDP – independent; IPFV – imperfective; NEG – negative; OBL – oblique; PL – plural; POSS – possessive; SG – singular; STAT – stative; SUBJ – intransitive subject; VF – volunteered form.

I analyze the stativizer as deriving a target state from a CoS root by existentially closing an event variable in the root, and foregrounding an underlying stative variable, essentially following Kratzer's (2000) analysis of derived statives in German. I discuss several properties of CoS roots and statives in Nsyilxcn which together support the idea that CoS root templates contain both event and state arguments. This analysis of Nsyilxcn CoS roots is more in line with Kratzer's (2000) analysis of underived German participle stems, and contrasts with recent analyses of English CoS roots where root templates come prespecified with either an existentially closed event (Beavers & Koontz-Garboden 2020) or state argument (Yu et al. 2023), as well as Davis' (2021) analysis of CoS roots in the neighbouring Salish language St'át'imcets, which lack state arguments altogether.

2 Target states and resultant states

As originally described in Parsons (1990), and formalized in Kratzer (2000), *target states* are in principle reversible, and the state must continue to affect an argument relative to a reference time in order to be felicitously used. *Resultant states*, in contrast, simply entail that an event has culminated at some point prior to the reference time (like the English perfect), and as such are not reversible, and do not require the state to continue affecting an argument at a reference time. One empirical test distinguishing the two types of states involves the adverb *still*, which occurs with target states like *pumped up* (2a), but is redundant with resultant states like *proven* (2b).²

- (2) a. The tires are *still* pumped up. [TARGET STATE]
 b. The theorem is (**still*) proven. [RESULTANT STATE]
 (Kratzer 2000:385-386)

Davis et al. (2020) developed a series of storyboards (Burton & Matthewson 2015) designed to test whether derived statives in two Salish languages, St'át'imcets and ʔayʔajuθəm, denote *target* states or *resultant* states. Their test results indicate that while the St'át'imcets stativizer *es-*derives a resultant state, ʔayʔajuθəm stative reduplication derives a target state. Given that variation within the Salish family exists, it is important to determine how Nsyilxcn (ə)c- statives pattern.

² The results of the *still* test are clearer in German than they are in English, as noted by Embick (2009).



Figure 1: The Broken Cup



Figure 2: The Breakdown



Figure 3: The Trodden Worm

Figure 1 represents the final pane of a storyboard about a woman who drops a cup, whose broken pieces are scattered, after which she uses glue to piece the cup back together. The stative *cpak^w* ‘to be scattered’ is volunteered earlier in the storyboard to describe a pane in which the shattered pieces lay strewn about the floor. If *cpak^w* denotes a resultant state, it should be felicitous to use even after the cup has been glued back together (Figure 1), similarly to the English present perfect *It has been scattered (but is now put back together)*. If *cpak^w* denotes a target state, it should not be felicitous in Figure 1, since the state no longer actively affects the cup. Results indicate a target state (3).³

- (3) #ʕapnáʔ **c-pak^w** iʔ lpot.
 now STAT-scattered DET cup
 ‘The cup has now been scattered.’

(Delphine Derickson Armstrong)

The final scene in the second storyboard is illustrated by Figure 2. This tells the story of a couple whose car breaks down, after which they try to push the car to a service station. For one version, they successfully push the car, and the stative *cyrmin* ‘to be pushed’ is volunteered.⁴ In an alternate version (Figure 2), the couple tries to push the car, but to no avail. Under this scenario, a pushing event has occurred, but there is no discernable target state affecting the car. Hence, the prediction is that

³ Note that the adverb *ʕapnáʔ* ‘now’ is necessary to enforce a present tense reading of the stative. Without *ʕapnáʔ*, the sentence *cpak^w iʔ lpot* is preferably interpreted relative to a past time, before the cup is glued together again, i.e. *The cup was scattered*, but because in this case, the stative is interpretable either as a resultant or as a target state, the test is invalidated. Similar facts hold for (4) below, but in (5b) Delphine’s comments indicate that she is assigning a present tense interpretation even in the absence of *ʕapnáʔ*.

⁴ *yrmín* ‘to get pushed’ has the morphological appearance of an applicative, rather than a typical CVC root. The applicative *-min* seems to be historically fused, however. The etymological root, *y(i)r*, refers to ‘circling’.

The representation in (8) states that relative to a world w , an individual x undergoes a transitional P event e (via BECOME, Dowty 1979), and this event causes a state s (via CAUSE, Embick 2009).⁷

3 State variables and causative structure in change-of-state roots

In this section I provide evidence that a semantically causative predicate (CAUSE in 8) linking a transitional event with a state is an important component of meaning in Nsyilxcn CoS roots.

First, consider that there is a homophonous, yet semantically distinct imperfective prefix $(ə)c-$. Stage-level adjectives but not individual-level adjectives (Carlson 1977; Kratzer 1989) can occur with imperfective $(ə)c-$ (Lyon 2010). The contrast between I-level (9a) and S-level (9b) follows if the imperfective requires a predicate with an open event variable, and if S-level but not I-level adjectives have such a variable.

- (9) a. **(*c)-tíkʷəlqʷ** $iʔ$ sqəltmíxʷ.
 IPFV-tall DET man
 ‘The man is tall.’ (Delphine Derickson Armstrong)
- b. axáʔ **(c)-nʃast** t knəxnáx.
 this IPFV-heavy OBL box
 ‘This is a heavy box.’ (Delphine Derickson Armstrong)

Although historically related to stative $(ə)c-$, the imperfective is distinct: This is shown by the presence of habitual readings with imperfective adjectives (10), and their conspicuous absence on derived statives (11,12).⁸ In other words, imperfective $(ə)c-$ does not occur with bare CoS roots.

⁷ Evidence from manner adverbs and instruments of causation discussed below suggests that in at least some cases CoS roots combine with their internal arguments prior to stativization, yielding a *phrasal* stativization (cf. Kratzer 2000:7).

⁸ There is nothing inherent about the determiner $iʔ$ that should force reference to a single entity in (11) and (12): $iʔ$ allows generic readings (Lyon 2015), but stative $c-$ seems to prevent a generic interpretation. The habitual readings targeted in (11) and (12) are felicitously expressed using a transitive, causative imperfective, e.g., *cpʷqstixʷ iʔ stiqʷ* ‘You (typically) cook meat’ and *cpulstəlx iʔ siʔyʷ* ‘They (typically) tanned hides’. This shows that CoS roots can be derived into forms which are compatible with the imperfective, and suggests that the causativizer *-st-* may be interacting semantically with the state variable in a CoS root.

- (10) uc k^w c-ʔilx^wt?
 DUB 2SG.SUBJ IPFV-hungry
 ‘Don’t you get hungry (typically)?’ (twi-Lottie Lindley)
- (11) *Context: Showing someone new around a kitchen.*
 #c-ḡyq iʔ sʔiq^w aláʔ iʔ l nk^wícncútən.
 STAT-get.cooked DET meat here DET in cooking.container
Target: ‘Meat is cooked in this pot.’
Actual: ‘The meat has been cooked in this pot.’
 (Delphine Derickson Armstrong)
- (12) #q̄sápi c-puí iʔ sip̄ȳ.
 long.ago STAT-get.tanned DET hide
Target: ‘Long ago, hides were tanned.’
Actual: ‘Long ago, the hide was tanned.’
Comment: ‘You’re just talking about one hide.’
 (Delphine Derickson Armstrong)

If S-level adjectives and CoS roots were both simply predicates over eventualities, the expectation is that habitual, imperfective interpretations of CoS roots should be possible in (11) and (12), as they are with a wide range of derived verbal predicates. Instead, only a stative reading obtains. Assuming (i) an analysis of imperfective (ə)c- similar to that in (13) (see Rullmann & Matthewson 2018), and (ii) that CoS roots contain an additional, open state argument necessary for deriving target statives, the prediction is that imperfective c- may not occur with CoS roots for compositional reasons: after saturation of the internal argument they are of type ⟨s,⟨s,t⟩⟩ (8) rather than ⟨s,t⟩, as required by (13).⁹

$$(13) \quad [[c\text{-IPFV}]] = \lambda P_{\langle s,t \rangle} \lambda t \exists e. [P(e) \wedge t \subseteq \tau(e)] \quad [\text{IMPERFECTIVE}]$$

This general approach is supported by an additional fact: while adjectives commonly occur as bare unaccusatives in Nsyilxcn (9), CoS roots never occur as bare unaccusatives (14).¹⁰ The reason for this, I suggest, is that having both stative and eventive arguments open, CoS

⁹ This approach also presumably rules out (null) perfective, or neutral (Smith 1991), interpretations of bare CoS roots.

¹⁰ This is in stark contrast to CoS roots in St’át’imcets (Lyon & Davis 2022). Lyon (2023) shows how agentive uses of bare CoS roots in Nsyilxcn are analyzable as zero-derived middles.

Note that the adverbial phrase *t x^wúsx^wəst* ‘quickly’ in (15b) must be interpreted as modifying the event of being stolen. In some cases, however, a manner adverbial seems forced into infelicitously modifying a state argument (16).

- (16) ***c-čax^w** i? siwłk^w t kəkali?.
 STAT-get.spilled DET water OBL slow
 ‘The water is spilled slowly.’
Comment: “How can it be *kəkali?* when it is already spilled?!”
 (Delphine Derickson Armstrong)

The contrast between (15b) and (16) hints that manner adverbs can attach in different locations syntactically, and that this has interpretive implications: The event-modifying adverb in (15b) attaches before the event variable is existentially closed by the stativizer (cf. 7a), while the adverb in (16) is interpreted as attaching afterwards.¹²

Lastly, homogenous S-level adjectives cannot host an oblique instrument of causation (17a), while derived statives can (17b).

- (17) a. ***łəst[?]** i? lasmíst i? t sqit.
 wet DET shirt DET OBL rain
 ‘The shirt was made wet by the rain.’
 (Delphine Derickson Armstrong)
- b. way **c-nik[?]** i? spícən i? t krkriwstn.
 already STAT-get.cut DET rope DET OBL scissor
 ‘The rope was cut by the scissors.’
 (Delphine Derickson Armstrong)

Oblique instruments require reference to a causing event (Davis & Demirdache 2000), i.e. *CAUSE(e,s)* in (8) (Embick 2009), and by extension a change-of-state.¹³ The grammatical patterns in (15) to (17) follow if adjectives do not encode any change-of-state or causing event,

¹² It is currently unclear what contextual factor(s) determine whether a speaker interprets a manner adverb as modifying an event (15b) or a state (16): the important point for now is that either interpretation is, in principle, possible.

¹³ This argument receives empirical support from examples involving inchoativized predicates, which I analyze as predicates over events. While inchoatives formed from CoS roots allow modification of an underlying state, inchoatives formed from adjectival roots do not (Lyon, 2024).

while derived stative forms contain both eventive and stative arguments, linked together by a *CAUSE* predicate.

If a saturated event argument, closed by the stativizer in (7), is what forces an illicit manner adverbial modification of a result state in cases like (16), this implies that the manner adverb in (15b) and the oblique instrument in (17b) must be referencing the event *prior* to that event argument being saturated, i.e. *prior* to stativization. From this, it follows that the change-of-state and causing event must be present in the CoS root itself, rather than being introduced by the stativizer to a simplex predicate over events (Embick 2009).

4 Discussion

Nsyilxcn provides support for Kratzer's (2000) definition of the target stativizer in German, and the distribution of stativizer *(a)c-* helps provide evidence for a semantic distinction between lexical classes in Nsyilxcn: verbal CoS roots are semantically causative (Davis & Demirdache 2000), while adjectives are not. Nsyilxcn additionally shows that it is possible that English CoS roots might be amenable to a more abstract analysis than that recently proffered by Yu et al. (2023) or Beavers and Koontz-Garboden (2020): If English CoS roots can be analyzed similarly to underspecified Nsyilxcn CoS roots or underived German target state participle stems (Kratzer 2000), then they may have zero derivations into stative and eventive forms, supporting an analysis similar to Lieber (1980) who proposes that English and German adjectival participles contain a zero-stativizer.

This paper also shows that Nsyilxcn is different from other Salish languages such as St'át'imcets (Davis 2021; Lyon & Davis 2022) in that unaccusative CoS roots may not be used in bare form. The reason for this, I have suggested above, is that CoS roots are pre-equipped with open event *and* state variables, and for this reason are underspecified without further derivation. Stativizer *c-* backgrounds the event argument (by existential closure of the *e* variable), and foregrounds the resulting state (Kratzer 2000; Burton & Davis 1996 for St'át'imcets) which resolves the underspecification issue, leaving only the stative portion open for temporal modification. This means that a simplex eventive analysis of CoS roots, similar to that advanced in Davis (2021) for St'át'imcets CoS roots, will not suffice for Nsyilxcn. Davis (2021) may nevertheless be correct about St'át'imcets CoS roots, considering that these derive into resultant states. Assuming that Nsyilxcn and St'át'imcets CoS roots, though both unaccusative, differ semantically in whether they contain an

underlying state variable and causative semantics, the conclusion is that the Unaccusativity Hypothesis for Salish (Davis 1997) must accommodate some degree of variation.

Finally, there are two historical points worth making. First, if the Nsyilxcn imperfective $(ə)c-$ has its origins as a stativizer, we might expect some semantic similarity between the two markers, especially if the divergence is somewhat recent. Target states and imperfectives both share a requirement that an eventuality be in the process of affecting an argument relative to a reference time, and both of these contrast with resultant states in this respect. As such, this analysis accords with a common historical root for the two $c-$ prefixes. Second, given the cognacy between resultant state-deriving St'át'imcets $es-$, Secwepemctsin $s-/c-$ (Kuipers 1974), and target state-deriving Nsyilxcn $(ə)c-$, it is possible that Proto-Nsyilxcn $*(ə)c-$ shifted from deriving a resultant state to deriving a target state, and that this conditioned the use of $(ə)c-$ as an imperfective marker. At the same time, it is possible that causative event structure and a state variable originally contributed by Proto-Nsyilxcn $*(ə)c-$ to simplex eventive CoS roots may over time have become reanalyzed as part of the lexical meaning of CoS roots in Nsyilxcn, accounting for the difference between CoS roots in St'át'imcets and Nsyilxcn: Since target states have both an event and a state variable, a shift in the semantics of the stativizer from resultant state-denoting to target state-denoting would require a concomitant shift in the semantics of CoS roots. Alternatively (and equivalently in terms of its semantic effect) a reanalysis of CoS roots as containing a stative argument may have forced a shift in the semantics of the stativizer. Further work on statives across Salish may help to illuminate historical connections between resultant and target states.

5 Conclusion

This paper helps address some of the gaps in language documentation relating to lexical aspect (A. Mattina 1993; N. Mattina 1996), with the aim of establishing a root-level semantics to provide a firm basis for further work. I show how the Nsyilxcn stativizer $(ə)c-$ derives an unaccusative target state (Kratzer 2000; Davis et al. 2020), and provide evidence that change-of-state roots contain both stative and event arguments. This research complements previous aspectual studies for Salish languages (Bar-el 2005; Kiyota 2008; Davis et al. 2020), provides insight into the event structure of verbal roots, raises interesting questions regarding possible semantic variation across Salish at the root

level, and has implications for semantic theories of lexical roots and how they relate to event structure (Kratzer 2000; Embick 2009; Beavers & Koontz-Garboden 2020; Yu et al. 2023).

References

- Bar-el, Leora. 2005. *Aspectual Distinctions in Skwxwú7mesh*. Ph.D. Dissertation. University of British Columbia, Vancouver, BC.
- Beavers, John & Andrew Koontz-Garboden. 2020. *The roots of verbal meaning*. Oxford, UK: Oxford University Press.
- Burton, Strang & Henry Davis. 1996. Stative Aspect and Possession in Salish. *Papers for ICSNL 31*: 13–22.
- Burton, Strang & Lisa Matthewson 2015. Targeted construction storyboards in semantic fieldwork. In Ryan Bochnak & Lisa Matthewson (eds.), *Semantic fieldwork methodology*. Oxford, UK: Oxford University Press, 135–156.
- Carlson, Gregory. 1977. *Reference to Kinds in English*. Ph.D. dissertation, University of Massachusetts, Amherst, MA.
- Davis, Henry. 1997. Deep unaccusativity and zero syntax in St’át’imcets. In Andrea Mendikoetxea & Myriam Uribe-Etxebarria (eds.), *Theoretical Issues at the Morphology-Syntax Interface*. Bilbao, Spain: Supplements of the International Journal of Basque Linguistics and Philology, 55–96.
- Davis, Henry, Marianne Huijsmans, & Gloria Mellesmoen. 2020. Statives in ʔayʔajuθəm and St’át’imcets. *Papers for ICSNL 55*:107–132.
- Davis, Henry. 2021. Going Radical in Salish. *Proceedings of the 2021 West Coast Conference on Formal Linguistics (WCCFL 39)*.
- Davis, H. & H. Demirdache. 2000. On Lexical Verb Meanings: Evidence from Salish. In Carol Tenny & James Pustejovsky (eds.), *Events as Grammatical Objects. The Converging Perspectives of Lexical Semantics and Syntax*. Stanford, CA: CSLI Publications, 97–142.
- Dowty, David. 1979. *Word Meaning and Montague Grammar: The Semantics of Verbs and Times in Generative Semantics and in Montague’s PTQ*. Dordrecht, Netherlands: Reidel.

- Embick, David. 2009. *Roots, states, stative passives*. Presented at the 2009 Roots Workshop, University of Stuttgart.
- First Peoples' Cultural Council. 2022. *The Report on the Status of B.C. First Nations Languages*. Brentwood Bay, BC.
- Kiyota, Masaru. 2008. *Situation Aspect and Viewpoint Aspect: From Salish to Japanese*. Ph.D. dissertation, University of British Columbia, Vancouver, BC.
- Kratzer, Angelika. 1989. Stage-level and individual-level predicates. *University of Massachusetts Occasional Papers in Linguistics* 15(2): Article 10.
- Kratzer, Angelika. 2000. Building statives. In Lisa J. Jonathan, Jeff Good, Darya Kavitskaya, Alyssa B. Wulf, & Alan C. L. Yu (eds.), *Proceedings of the Twenty-Sixth Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on Aspect*. Berkeley, CA: Berkeley Linguistics Society, 385–399.
- Kuipers, Aert. 1974. *The Shuswap Language: Grammar, Texts, Dictionary*. The Hague, Netherlands: Mouton.
- Lieber, R. 1980. *On the Organization of the Lexicon*. Ph.D. dissertation, MIT, Cambridge, MA.
- Lyon, John. 2010. *Nominal Modification in Okanagan: Complex DPs, Complex Nominal Predicates, and Stative/Customary ac-*. Presented at ICSNL 45 in Eugene, OR.
- Lyon, John. 2015. Okanagan Determiner Phrases and Domain Restriction. *International Journal of American Linguistics* 81(2):187–219.
- Lyon, John. 2023. Building Statives and Inchoatives in Nsyilxcn. *Papers for ICSNL 58*:230–290.
- Lyon, John. 2024. The Stative/Inchoative Alternation in Nsyilxcn. Manuscript under review.
- Lyon, John & Henry Davis (eds.). 2022. *Wa7 sqwéqwel' sSam: St'át'imcets Stories from Sam Mitchell*. Vancouver, BC: Pacific Northwest Languages & Literatures Press.
- Rullmann, Hotze & Lisa Matthewson. 2018. Towards a theory of modal-temporal interaction. *Language* 94(2):81–331.

- Mattina, Anthony. 1993. Okanagan aspect: A working paper. *Papers for ICSNL* 28:233–263.
- Mattina, Nancy. 1996. *Aspect and Category in Okanagan Word Formation*. Ph.D. dissertation, Simon Fraser University, Burnaby, BC.
- Parsons, Terence. 1990. *Events in the Semantics of English. A Study in Subatomic Semantics*. Cambridge, MA: MIT Press.
- Smith, Carlota. 1991. *The Parameter of Aspect*. Dordrecht, Netherlands: Kluwer.
- Yu, Jianrong, Josep Ausensi, & Ryan W. Smith. 2023. States and changes-of-state in the semantics of result roots: Evidence from resultative constructions. *Natural Language & Linguistic Theory* 41:1589–1628.

A corpus-based study of Gitksan modals¹

LISA MATTHEWSON

University of British Columbia

1 Introduction

In this paper I test the predictions of previous analyses of Gitksan (Tsimshianic) modals on a corpus of 36 Gitksan stories. See Reisinger et al. (2022) for a similar (but more in-depth) corpus-based study on modals in English and St'át'imcets (Lillooet Salish).

Section 2 gives some necessary theoretical background. Section 3 describes the methodology of the study and introduces the four modals to be investigated. Sections 4 to 7 summarize the findings for each modal, and Section 8 concludes.

2 Theoretical background

Modals are standardly analyzed as quantifiers over possible worlds (Kratzer 1991). I will be investigating two core properties of modals: their modal force, and their modal flavour. *Modal force* refers to the quantificational strength of the modal. Example (1) presents some English modals that lexically encode differing modal forces, from strongest in (1a) to weakest in (1c).

- (1) a. NECESSITY:
Zoe **must** meet with her thesis supervisor.
- b. WEAK NECESSITY:
Zoe **should** meet with her thesis supervisor.

¹ This paper is inspired by, and follows on from, collaborative work with Hotze Rullmann (Rullmann et al. 2008; Rullmann and Matthewson 2018; Reisinger et al. 2022). Hotze has been a great colleague, research collaborator, mutual supervisor of students, and friend for nearly 20 years so far. Much of what I have done in my career, I couldn't have done without Hotze, and I am very grateful.

Many thanks to Gitksan speakers Vincent Gogag, Hector Hill, and Barbara Sennott for their beautiful stories and for their work over many years documenting their language. *Ha'miyyaa!* Many thanks to the members of the Gitksan Research Lab who have contributed to the Forbes et al. (in prep.) volume — especially Clarissa Forbes, Michael Schwan, and Henry Davis, and many others over the years. Thanks also to Henry Davis and Clarissa Forbes for commenting on a draft of this paper.

- c. POSSIBILITY:
 Zoe **may** meet with her thesis supervisor.

Modal flavour refers to the type of modal reasoning that is involved; the different flavours result from restrictions on the sets of possible worlds that are quantified over (Kratzer 1991). Some flavours are illustrated for English in (2). Note that most or all non-epistemic flavours are sub-types of circumstantial modality. Thus, pure circumstantial, deontic, and ability flavours can all be grouped under circumstantial modality. This will become relevant below when we see the lexical distinctions that Gitksan modals make.

- (2) a. EPISTEMIC:
 Zoe **might** be in her office (her office door is open).
- b. PURE CIRCUMSTANTIAL:
 Roses **might** grow here (the soil and climate are right).²
- c. DEONTIC:
 Zoe **should** be in her office (according to the rules).
- d. ABILITY:
 Zoe **can** lift 50 kilos.

Languages differ in whether they tend to lexically encode modal force or modal flavour (or both, or neither). As seen in (1), English often lexically distinguishes modal force, and as seen in (2), English often does not lexically distinguish modal flavour. For example, the single lexical item *might* can be interpreted either epistemically or circumstantially.

Another important facet of modality is modal-temporal interactions (Condoravdi 2002). For space reasons, I focus here on only one aspect of these interactions, namely temporal orientation. This refers to whether the postulated event takes place before, simultaneously with, or after the time at which the modal is evaluated. These options are illustrated for English in (3). In all these examples, the modal is evaluated at the utterance time (i.e., based on utterance time knowledge). The postulated event either precedes, coincides with, or follows the utterance time (UT).

- (3) a. PAST TEMPORAL ORIENTATION:
 Zoe must have arrived by now. (arrive < UT)

² Example adapted from Kratzer (1991:646).

- b. PRESENT TEMPORAL ORIENTATION:
Zoe might be arriving now. (arrive = UT)
- c. FUTURE TEMPORAL ORIENTATION:
Zoe might arrive soon. (UT < arrive)

3 The study

The corpus for this study is Forbes et al. (in prep.), a volume of 36 stories told by three Gitksan speakers. The speakers are Vincent Gogag, from Git-anyaaw (Gitanyow), Hector Hill, from Gijigyukwhla (Gitsegukla), and Barbara Sennott, from Ansba'yaxw (Kispiox). The corpus contains a little over 12,500 Gitksan words, and the stories have been translated into English and fully morpheme-glossed. Each line was translated into English by the storyteller, so we have the original speaker’s English rendition of all the sentences.

The modals to be tested are listed in Table 1, with prior proposals about their flavour and force. The two modals classified as ‘circumstantial’ allow all circumstantial sub-flavours, including pure circumstantial, deontic, and ability. The epistemic modal *ima'(a)* is analyzed by Peterson (2010) as also conveying an evidential restriction; the speaker must have inferential evidence for the prejacent proposition. The only modal omitted from this study is *gat*, the reportative evidential (Peterson 2010; Matthewson 2013). This is for space reasons and also because there were only 14 tokens of *gat* in the corpus, all from one speaker and almost all from a single story.

Table 1: Gitksan modals

MODAL	FLAVOUR	FORCE	REFERENCES
<i>sgi</i>	circumstantial	(weak) necessity	Rigsby (1986); Matthewson (2013)
<i>da'akhlxw</i>	circumstantial	possibility	Rigsby (1986); Matthewson (2013)
<i>anook</i>	deontic	possibility	Rigsby (1986); Matthewson (2013)
<i>ima'(a)</i> ³	epistemic	variable force	Peterson (2010, 2012); Matthewson (2013)

³ This modal is pronounced as *imaa*, *ima'*, or *ima'a*, depending on the speaker and possibly on speech rate.

Regarding temporal orientation, Matthewson (2012, 2013) has argued that future orientation is always overtly spelled out in Gitksan via the prospective aspect marker *dim*, while past and present temporal orientation are not overtly encoded (see also Matthewson & Todorović 2018; Rullmann & Matthewson 2018). This predicts that the epistemic modal *ima('a)* will co-occur with *dim* when — and only when — the temporal orientation is future. It further predicts that all the circumstantial modals (*sgi*, *da'akhlxw*, and *anook*) will always co-occur with *dim*, since circumstantial modals are by their very nature future-oriented (see Condoravdi 2002; Werner 2006; however, see Thomas 2014 for the proposal that this only holds for pure circumstantials).

All instances of all the modals were identified by searching for them by gloss. This resulted in a total of 19 tokens of *sgi*, 34 tokens of *da'akhlxw*, 26 tokens of *anook*, and 32 tokens of *ima('a)*. The sentences the modals appeared in were inspected for their meaning, using both the translation into English and the surrounding context. Each token was coded using the categories in Table 2.

Table 2: Categories used in the annotation process

Categories	Annotation options
flavour	epistemic pure circumstantial deontic ability undetermined
force	necessity weak necessity possibility undetermined
temporal orientation	past present future undetermined

In the following sections I present the findings. This is not a statistical study; only qualitative comments plus some raw numbers will be presented.

4 *Sgi*

According to prior research, *sgi* should appear with exclusively circumstantial flavours; Matthewson (2013) establishes its use with pure circumstantial, deontic, and teleological flavours, and notes that one of the most common flavours of *sgi* is deontic (2013:380).⁴ Matthewson

⁴ Matthewson (2013:382–383) observes one gap: *sgi* does not allow strong necessity pure

The prediction that *sgi* will always co-occur with *dim* is supported insofar as 18 of the 19 *sgi* tokens have *dim* on the modal's preadjacent. The only exception is the sentence in (5).⁷

Sometimes deontic modality is conveyed in the English translation, but in Gitksan only the prospective *dim* is used; examples of this are given in (6) and (7).

- (6) 'Ni[t]=gan wil[-t]=hl wen-i'm di-ye dip
 3.III=reason be/do[-3.II]=CN sit.PL-1PL.II 3.I=QUOT ASSOC
 nibib-i'y, **dim**=in dip hlimoo-dit 'nuu'm, **dim**
 uncle-1SG.II **PROSP**=1.I 1PL.I help-3PL.II 1PL.III **PROSP**
 hehle'lsd-i'm loo-dit.
 work-1PL.II OBL-3PL.II
 'That's why my uncle says, he says we are to help them, we are to
 work for them.'

(Hector Hill, Hlaa yukw dim 'nu'whl get /
 Before the people die, line 83)

- (7) K'ap=hl gabi=hl ayook **dim** luu yuxw[-i]-diit.
 ten=CN how.many=CN law **PROSP** in follow[-TR]-3PL.II
 'There are ten laws that they should follow.'
 (Barbara Sennott, Dihlxw / The boat, line 11)

Matthewson et al. (2022) argue that when *dim* appears without a modal, there is a phonologically null modal in the structure above *dim*. This allows us to maintain a unified analysis of *dim* as providing only temporal ordering. Thus, *dim* always serves to provide the future orientation for some modal element. The analysis accounts for (6) and (7), but future research is required to establish whether there are force and flavour constraints on the null modal.

5 *Da'akhlxw*

Da'akhlxw has been argued to be a general circumstantial possibility

⁷ Clarissa Forbes (p.c.) suggests that the sentence-initial *dim* in (5) may have started out following *sgi*, in the position where we expect it to appear, and moved to the front. Forbes notes that the presence of the 1st person plural pronoun *dip* could somehow be responsible, as *dim* has been noticed in an unexpected sentence-initial position in other constructions with *dip*. Further research is required.

modal. This predicts that it will allow pure circumstantial possibility readings, deontic possibility readings (i.e., permission), and ability readings.⁸ We also predict that it will obligatorily co-occur with *dim*, like other circumstantial modals.

Of the 34 tokens of *da'akhlxw* in the corpus, 30 convey ability readings; one example is (8).

- (8) si-t'aa-'m-am sik'ihl huut-xw-dii, ii nee
 CAUS1-sit-DETR-ATTR try flee.PL-VAL-3PL.II CCNJ NEG
 dii helt[=hl] get ji an=t **da'akhlxw**[-t]
 FOC many[=CN] people IRR AX=3.I CIRC.POSS[-3.II]
 dim huut-dii.
 PROSP flee.PL-3PL.II
 'They tried to flee, but not many were able to flee.'
 (Vincent Gogag, *Xhluxwhl sga'nist go'ohl ksi txemsim /*
The Nass River volcano, line 25)

A further three tokens convey deontic possibility (permission), as illustrated in (9).

- (9) "Nee dii **da'akhlxw**[-t] dim ma<has>'us-in ji nee
 NEG FOC CIRC.POSS[-3.II] PROSP <PL>play-2SG.II IRR NEG
 mi dii sdil-i'm."
 2.I FOC accompany-1PL.II
 "“And you can't play around if you're not going to come with us.”"
 (Hector Hill, *Jayehli'm / Our traps*, line 18)

The last of the *da'akhlxw* tokens, given in (10), does not clearly portray circumstantial possibility. However, it is plausible that *da'akhlxw* is used here as a politeness device, much as in English we can say 'Can you tell us the story?'. Note that the use of negation in (10) is standard for an un-biased polar question in Gitksan; see Rigsby (1986:296); Matthewson (2022); Hill and Matthewson (in prep.).

⁸ The modal force of ability modals is actually a matter of debate in the literature, and many have observed that a standard possibility analysis is too weak (see e.g., Portner 2009:201–203, and Louie 2014:160ff. and references therein). However, most analyses of ability modals have an existential quantifier somewhere in the denotation, and the fact is that Gitksan lexically groups ability readings with readings that are uncontroversially possibility readings, such as permission.

- (10) “Nee=m **da'akhlxw**[-t] dim=a mehli[-t]=hl wila
 NEG=2.I CIRC.POSS[-3.II] PROSP=2.I tell.T[-3.II]=CN MANR
 wi[1][-t]=hl betl'-a betl' loo-'m=aa?”
 be/do[-3.II]=CN plop-ATTR plop OBL-1PL.II=Q
 “‘Will you tell us the story about *betl'a betl'*?’”⁹
 (Hector Hill, *Betl'a betl'* / Story of a name, line 39)

Of the 34 tokens of *da'akhlxw* in the corpus, 28 of them straightforwardly co-occur with a following *dim*, as predicted; this can be seen in examples (8) to (10). In a further four cases, there is no *dim*, but this is because the prejacents is completely elided; an example of this is given in (11).

- (11) 'Nidiit[=hl] dim waatxw-it a[-t]=hl get ji
 3PL.II[=CN] PROSP cry-SX PREP[-3.II]=CN person IRR
 nee ji[=t] **da'akhlxw**-diit.
 NEG IRR[=3.I] CIRC.POSS-3PL.II
 ‘They are the ones to cry when the people that lost a person can’t
 [cry].’
 (Hector Hill, *Hlaa yukw dim 'nu'whl get* /
 Before the people die, line 28)

The two remaining tokens that lack *dim* are given in (12) and (13). In (12), Clarissa Forbes observes (p.c.) that the material following *da'akhlxw* is not a clause, but a nominal (relative clause). *Da'akhlxw* here seems to have the meaning ‘manage to obtain (a thing)’, and may be a separate construction.

- (12) Iit dok[-t]=hl walk'a 'nit=hl gabii=hl
 CCNJ=3.I take.PL[-3.II]=CN all 3.III=CN how.many=CN
 dim hooy-i-t, dim wila=t **da'akhlxw**[-t]
 PROSP use-TR-3.II PROSP MANR=3.I CIRC.POSS[-3.II]
 siilinas-xw-t.
 hunt-ANTIP-3.II
 ‘And he gathered everything to use so he could catch what he’s
 hunting.’
 (Hector Hill, *Betl'a betl'* / Story of a name, line 6)

⁹ *Betl'a betl'* is a rendition of the noise a grouse makes when it flies.

- (13) *Ii=t da'akhlxw-diiit ksi sim-guu-t-diiit.*
 CCNJ=3.I CIRC.POSS-3PL.II out true-take-T-3PL.II
 'They were able to wrestle it out of the water.'
 (Vincent Gogag, Wilps Gu'nuu /
 The House of Gu'nuu, line 14)

As seen in (11), *da'akhlxw* can scope under negation, with the meaning 'not able to' or 'not allowed to'. The corpus also revealed two monomorphemic forms to express inability: *hlguxws* (in Barbara Sennett's stories) and *gos* (in Hector Hill's stories). The difference, if any, between these forms and negated *da'akhlxw* could be followed up in future research.¹⁰

6 *Anook*

Anook has been analyzed as a deontic possibility modal; the prediction is therefore that it should be used exclusively to convey permission interpretations. It should obligatorily co-occur with *dim*.

The predictions are upheld with near perfection: all 26 tokens of *anook* in the corpus have permission interpretations, and 25 of the 26 tokens either co-occur with *dim* (22 tokens), or appear with a fully elided preadjacent (three tokens). Examples are given in (14) and (15), with and without an overt preadjacent, respectively.

- (14) *Ii=t anook[-t]=hl sim'oogit dim=t 'nii*
 CCNJ=3.I DEON.POSS[-3.II]=CN chief PROSP=3.I on
t'aa-d-it goo=hl lax se'e-t.
sit-T-3.II LOC[-3.II]=CN on leg-3.II
 'And the chief allowed the stranger to have the baby sit on his lap.'
 (Barbara Sennott, Ha'niisgats 'Wii Gat /
 'Wii Gat's birth, line 20)

- (15) ... *ii da'akhlxw[-t]=hl dim=m sdil-i'm,*
 ... CCNJ CIRC.POSS[-3.II]=CN PROSP=2.I accompany-1PL.II
ji=da=t anook-diiit 'niin."
 IRR=SPT=3.I DEON.POSS-3PL.II 2SG.III
 '... and you can come with us, if they allow you to [come].'
 (Hector Hill, T'aahl isi'm / Picking soapberries, line 25)

¹⁰ Neither *hlguxws* nor *gos* are found in Rigsby (1986) or Hindle and Rigsby (1973). Tarpent (1987:485) mentions *gos* in Nisga'a and glosses it as 'can't do something'.

The one counter-example to the presence of *dim* with *anook* is given in (16). It is possible that this example is different because it has a negated preajcent, but further research is required.

- (16) *Ii=t anook-dii nee ho_x dii yee-'y.*
 CCNJ=3.I **DEON.POSS-3PL.II** NEG again FOC go-1SG.II
 ‘And they allowed me not to go.’
 (Hector Hill, Jayeehli'm / Our traps, line 66)

For completeness, I note that there are no tokens of *anook* in the corpus from Vincent Gogag. This is surely a coincidence (i.e., a result of the particular stories Vincent happened to tell here), as Matthewson (2013) provides multiple examples of *anook* collected from Vincent.

7 *Ima('a)*

Our final modal, epistemic *ima('a)*, provided the most surprising results.

The predictions for *ima('a)* are that it will have exclusively epistemic modal flavour, it will be compatible with any modal force, and its preajcent will contain *dim* when, and only when, the modal has future temporal orientation (i.e., when the hypothesized event takes place after the modal's temporal perspective).

There are 32 tokens of *ima('a)* in the corpus, and all of them seem to have epistemic modal flavour. An example is given in (17).

- (17) *Sib-in[-i-t]=hl sim'oogit=hl dilhxw. Sga*
 hard-CAUS2[-TR-3.II]=CN chief=CN bag blocking.way
 ts'iib-i-t=**imaa**.
 tie-TR-3.II=**EPIS**
 ‘The chief tightened up the boat. Perhaps he tied it off.’
 (Barbara Sennott, Dihlxw / The boat, line 49)

The other prediction that is straightforwardly upheld is the one about temporal orientation. The only three tokens where *dim* appears on *ima('a)*'s preajcent are cases of future temporal orientation, as illustrated in (18) (and also in (25) below):

- (18) Ligi 'negw=**ima'a** dim k'uhl wil-t.
 DWID long=**EPIS** PROSP around be/do-3.II
 “‘He may take a long time.’”
 (Vincent Gogag, Sga'watxw liksgigedim get /
 Adventures with strangers, line 22)

Modal force is difficult to determine for *ima('a)* in the corpus. Based on translation, approximately half the tokens can be classified with some confidence as having existential force (conveying possibility); (17) and (18) are examples of this, and (19) is another. Example (19) clearly involves a possibility interpretation, given the immediately preceding clause that expresses unsureness.

- (19) Nee dii=n wilaax[-t=hl] wila wil, ligi
 NEG FOC=1.I know[-3.II=CN] MANR be/do DWID
 sga hit'-in[i]-d=**imaa**.
 blocking.way stick-CAUS2[-TR]-3.II=**EPIS**
 ‘I don’t know how it is done, maybe sealed.’
 (Barbara Sennott, Dihlxw / The boat, line 50)

No tokens of *ima('a)* are translated with strong necessity modals (*must* or *have to*). There is one that is translated with ‘probably’, and may therefore have a weak necessity interpretation:

- (20) Ii he-diiit loo-t k'ap am gilbil[=hl] wilaax[-i]-
 CCNJ say-3PL.II OBL-3.II VER only two[=CN] know[-TR]-
 diid=**ima'a** a[-t]=hl k'amksiwaa-mx-diiit.
 3PL.II=**EPIS** PREP[-3.II]=CN white.people-language-3PL.II
 ‘They said that they only knew probably two words in English.’
 (Vincent Gogag, Sga'watxw liksgigedim get /
 Adventures with strangers, line 12)

The remainder of the tokens of *ima('a)* receive a range of translations. Some are translated with expressions of vagueness like ‘kind of’ or ‘about’, as in (21); these make sense, since saying there were ‘about four’ is similar to saying there were ‘maybe four’.

- (21) *li sagayt t̥xalpx=uma'a=hl gabi-'m saa*
 CCNJ together four=EPIS=CN how.many-1PL.II away/off
bax xba hlo'o-t lax sga'nist
 uphill mid go.PL-SX on mountain
 ‘And there was about four of us that walked up the mountain’
 (Hector Hill, Jayeehli'm / Our traps, line 35)

Another set of *ima('a)* tokens are translated with expressions like ‘I’m not sure’ or ‘I don’t know’, as in (22) and (23). Example (22) is literally ‘He might not come back’, and (23) is literally something like ‘It might be that many times that they took him around’.

- (22) *Ligi neey=ima'a dim dii gukws 'witxw-t,' d=iya.*
 DWID NEG=EPIS PROSP FOC back arrive-3.II 3.I=QUOT
 ‘‘‘We don’t know when he’ll be back.’’’
 (Vincent Gogag, *Sga'watxw liksgigedim get /*
Adventures with strangers, line 23)
- (23) *li day=imaa=hl gabii=t luu-tk'u di-yee-t.*
 CCNJ SPT=EPIS=CN how.many=3.I in-circular COM-go-3.II
 ‘And I’m not sure how many times they took him around.’
 (Barbara Sennott, *Bitxw / Divorce*, line 8)

Something that had not been noticed in prior literature is the frequent use of *ima('a)* in the formation of ignorance free relatives, as illustrated in (24) and (25).

- (24) *gan=hl aaty-asxw gan=hl, gwiy=imaa=t*
 PCNJ=CN feel-ANTIP PCNJ=CN what=EPIS=3.I
si-wad[-t]=ihl amxsiwaa.
 CAUS1-name.T[-3.II]=CN white.people
 ‘and spiritual visions and, whatever the white people call it.’
 (Barbara Sennott, *Gwiis gan'mala / Button blanket*, line 6)
- (25) *ii-t jap[-t]=hl gwiy=imaa dim hooy-i-t*
 CCNJ=3.I make[-3.II]=CN what=EPIS PROSP use-TR-3.II
dim=t jagw[-t]=ihl lalt.
 PROSP=3.I kill.T[-3.II]=CN snake
 ‘and he made whatever he was going to use to kill the snake.’
 (Barbara Sennott, *Wii lalt | Big snake*, line 11)

Future research is needed to work out the compositional semantics of these structures. Some but not all analyses of ignorance free relatives invoke epistemic modality, so the presence of *ima('a)* has the potential to make an interesting contribution here.¹¹

Occasional tokens of *ima('a)* are not translated at all, as in (26). This was also found in Reisinger et al.'s (2022) corpus-based study to be a feature of the St'át'imcets epistemic modal *k'a*.

- (26) T_xalpx=**uma'a**[=hl] gabii=hl aloohl bisde'y[=hl] gukws
 four=**EPIS**[=CN] how.many=CN INTJ grouse[=CN] back
 da-'witxw-i[-t]=s nigwood-i'm.
 COM-arrive-TR[-3.II]=PN father-1PL.II
 'Dad brought back four grouse.'
 (Hector Hill, Jayeehli'm / Our traps, line 60)

Finally, there are two cases where it is not obvious how to reconcile the speaker's English translation with the analysis of *ima('a)* as an epistemic modal. These are given in (27) and (28). In (27), there is a syntactic issue as well. According to Peterson (2010:70), "the most common surface position for =*ima* is as an enclitic to the first syntactic phrase in a clause." Peterson does note that there is some variability in the placement of *ima('a)*, but he does not give examples like (27), where the modal appears to attach to a sub-sentential constituent. The intended meaning seems to be that the speaker distances themselves from the name 'Indian Agent', but it is not immediately obvious how to derive this meaning from the extant analysis of *ima('a)* as an epistemic modal.

- (27) Way ii nee dii he[-t]=hl Indian Agent=**ima'a**,
 so CCNJ NEG FOC say[-3.II]=CN Indian Agent=**EPIS**
 si-wa-txws "Indian Agent", an=t
 CAUS1-name-VAL Indian Agent AX=3.I
 saayt-good-in[-t]=hl get,
 together-all.gone-CAUS2[-3.II]=CN people
 saayt-wen-dii.
 together-sit.PL-3PL.II
 'The Indian Agent disagreed, the so-called 'Indian Agent' who
 gathered the people together for the meeting.'
 (Vincent Gogag, Kitwancool reserve surveyed, line 15)

¹¹ See Šimík (2018) for a recent cross-linguistic discussion of free relatives, and references therein.

Example (28) is also interesting. Literally it seems to mean ‘It might have been recorded ...’, but it is translated as ‘It wasn’t recorded ...’.

- (28) Ligi t'imis=**ima'a**=hl k'uuhl luu-wen-diiit.
 DWID write=**EPIS**=CN year in-sit.PL-3PL.II
 ‘It wasn’t recorded how many years they were incarcerated.’
 (Vincent Gogag, Kitwancool reserve surveyed, line 41)

I have nothing to suggest about this apparent translation mismatch, but the presence of the element *ligi* in this example is worth mentioning. *Ligi* is a mysterious element that has not yet been formally analyzed. It has a range of uses: it appears in some free relatives, in free choice contexts, in disjunctions, on polarity indefinites, in combination with another element *wihl* ‘like’ to convey indirect evidentiality, and by itself to convey either ‘maybe’ or a vagueness/‘about’ interpretation. It is noteworthy that over a third of the *ima'a* tokens in the corpus — 11 out of 32 — contain *ligi*. (Apart from example (28), examples (18), (19), and (22) above contain *ligi*.) This is different from the other modals investigated in this study, which *never* appear with *ligi* unless it is otherwise required inside the preadjacent of the modal. This seems to suggest that *ligi* is somehow related to epistemic modality. Further research is definitely required.

8 Concluding remarks

Corpus-based studies of modals can provide a useful follow-up to elicitation-based studies. This small corpus-based investigation largely confirmed the predictions of prior research that had been based on hypothesis-driven fieldwork. The present study also revealed some things that had not emerged from that prior work. The most important of these relate to the epistemic modal *ima'a*, which is used in the corpus in a wider range of contexts than was expected. The corpus investigation also made clear that we need to get a handle on the connection and interplay between *ima'a* and the mysterious element *ligi*. It had not even come to my attention in my prior work on *ima'a* that it so frequently appears with *ligi*; Peterson (2010) also does not mention *ligi* in his work on *ima'a*.

At the same time, a corpus-based study alone would also not have revealed the full landscape. For example, we know from elicitation that the circumstantial necessity modal *sgi* has non-deontic uses, but the

corpus contains only deontic uses of *sgi*. Thus, we again see confirmed the truism that applying multiple data-collection methods leads to a fuller and more robust empirical picture.

References

- Condoravdi, Cleo. 2002. Temporal interpretation of modals: modals for the present and for the past. In David Beaver, Stefan Kaufmann, Brady Clark, & Luis Casillas (eds.), *The construction of meaning*. Stanford, CA: CSLI Publications, 59–88.
- Forbes, Clarissa, Henry Davis, Michael Schwan, & the UBC Gitksan Research Lab. (in prep.). *'Nit ganwilt (that's how it was): thirty-six Gitksan stories*.
- Hill, Hector & Lisa Matthewson. (in prep.). Aren't Gitksan polar questions biased? Ms., University of British Columbia.
- Hindle, Lonnie & Bruce Rigsby. 1973. A short practical dictionary of the Gitksan language. *Northwest Anthropological Research Notes* 7(1):1–60.
- Kratzer, Angelika. 1991. Modality. In Dieter Wunderlich & Arnim von Stechow (eds.), *Semantics: An international handbook of contemporary research*. Berlin, Germany: De Gruyter Mouton, 639–650.
- Louie, Meagan. 2014. *The temporal semantics of actions and circumstance in Blackfoot*. (Unpublished doctoral dissertation). University of British Columbia, Vancouver, BC.
- Matthewson, Lisa. 2012. On the (non-)future orientation of modals. In Ana Aguilar Guevara, Anna Chernilovskaya, & Rick Nouwen (eds.), *Proceedings of Sinn und Bedeutung* 16(2):431–446.
- Matthewson, Lisa. 2013. Gitksan modals. *International Journal of American Linguistics* 79:349–394.
- Matthewson, Lisa. 2022. Gitksan polar questions: negative or not? (Presentation). University of Chicago, Chicago.
- Matthewson, Lisa & Neda Todorović. 2018. Temporal properties of attitude complements: the Serbian – Gitksan connection. In Kimberly Johnson & Alexander Göbel (eds.), *Proceedings of SULA 10: The semantics of under-represented languages in the*

Americas, 149–163.

Matthewson, Lisa, Neda Todorović, & Michael Schwan. 2022. Future time reference and viewpoint aspect: evidence from Gitksan. *Glossa: A Journal of General Linguistics* 7(1):1–37.

Peterson, Tyler. 2010. *Epistemic modality and evidentiality in Gitksan at the semantics-pragmatics interface*. (Unpublished doctoral dissertation). University of British Columbia, Vancouver, BC.

Peterson, Tyler. 2011. The role of the ordering source in Gitksan modals. Paper presented at Semantics of Under-represented Languages in the Americas 6, University of Manchester.

Peterson, Tyler. 2012. The role of the ordering source in Gitksan epistemic modals. In Elizabeth Bogal-Allbritten (ed.), *Proceedings of SULA 6: The semantics of under-represented languages in the Americas*.

Portner, Paul. 2009. *Modality*. Oxford, UK: Oxford University Press.

Reisinger, Daniel, Lisa Matthewson, & Hotze Rullmann. 2022. Using corpora to investigate modal-temporal interactions. In Jozina Vander Kloek, Núbia Ferreira Rech, & Simone Guessier (eds.), *Modality in under-described languages: methods and insights*. Berlin, Germany: De Gruyter Mouton, 141–190.

Rigsby, Bruce. 1986. Gitksan grammar. Ms., University of Queensland, Australia.

Rullmann, Hotze, Lisa Matthewson, & Henry Davis. 2008. Modals as distributive indefinites. *Natural Language Semantics* 16:317–357.

Rullmann, Hotze & Lisa Matthewson. 2018. Towards a theory of modal-temporal interaction. *Language* 94(2):281–331.

Šimík, Radek. 2018. Ever free relatives crosslinguistically. In Uli Sauerland & Stephanie Solt (eds.), *Proceedings of Sinn und Bedeutung* 22(2):375–392.

Tarpen, Marie-Lucie. 1987. *A grammar of the Nisgha language*. (Unpublished doctoral dissertation). University of Victoria, Victoria, BC.

Thomas, Guillaume. 2014. Circumstantial modality and the diversity condition. In Urtzi Etxeberria, Anamaria Fălăuș, Aritz Irurtzun, &

Bryan Leferman (eds.), *Proceedings of Sinn und Bedeutung* 18:433–450.

Werner, Tom. 2006. Future and non-future modal sentences. *Natural Language Semantics* 14:235–255.

The Definiteness of Manners and Reasons*

MARCIN MORZYCKI
University of British Columbia

1 Introduction

As social scientists and airplane crash investigators must occasionally point out, almost nothing happens for only one reason. No plane ever crashed *just* because it was snowing or *just* because the pilot forgot to de-ice the wings, for example. It's only the conjunction of such factors that can provide an explanation. Fittingly, it's a mild insult to describe an explanation as 'monocausal'. That's because we recognize that the search for sole causes is naïve and often necessarily unrequited.

Despite all this, the linguistic choices we make often suggest sole causes, and it would be pedantic to object to that wording:

- (1) a. What is the reason Floyd left?
- b. The reason for Floyd's departure was a menacing swarm of bees.
- c. The reason for Floyd's departure is $\left\{ \begin{array}{l} \text{that} \\ \text{?because} \end{array} \right\}$ he was pursued by a menacing swarm of bees.

Superficially, the definite description *the reason* would seem to require a sole cause because definite descriptions generally require a unique referent. But even if Floyd is pursued by bees, he doesn't have *only* one reason to leave. He may have as many reasons as there are bees. He is also leaving because of not just the whole swarm but also because of various sub-pluralities of bees that make up the swarm. After all, a single menacing bee can be enough to trigger retreat. The point isn't trivial. The sentence is about what one might call a maximal reason, and that maximality has to come from somewhere. The only definite description

*This squib owes a debt to two Hotzes. One is the long-ago Hotze of the 1990s, whose dissertation shaped how we (for some suitably expansive value of 'we') think about maximality. The other is the Hotze of the present, who was been a fantastic colleague and much valued interlocutor about a large number of topics—including, most recently, how and why questions, which of course question manners and reasons. Hence this squib.

present is *the reason*, and it's headed by a singular count noun. But the maximal interpretation of definite descriptions normally arises only with plural and mass nouns. Even setting this issue aside, there must be other reasons for Floyd's departure, ones unrelated to the part structure of the swarm. At the risk of blaming the victim, he may well have *done* something to invite the wrath of the bees—perhaps poking their hive, say—and that too is a reason for his departure. Had it not happened, there would be no swarm.

The generalization seems to be that one can refer to *the reason* for an event without giving rise to the entailment that there is only one reason. That requires explanation.

It's not just reasons that work this way. Manners do too.¹ For example, just as one might ask about the reason for something, one can also ask for the manner in which it was done:

- (2) a. What is the way (in which) Floyd ran?
 b. The way (in which) Floyd ran was $\left. \begin{array}{l} \text{by taking huge strides} \\ \text{on his tiptoes} \\ \text{?quickly} \\ \text{?as fast as he could go} \end{array} \right\}$.

There is some syntactic awkwardness around putting an adverbial in predicate position in (2), but the general shape of the puzzle is the same as in (1). If Floyd ran by taking huge strides, he may have also run ridiculously or awkwardly, and saying that he ran in one of these ways doesn't give rise to the inference that he didn't also run in the other ways as well.

This squib suggests a way of looking at these facts. Section 2 illustrates the effect more fully, focusing on paradoxical behavior with respect to cardinality. Section 3 observes similar behavior in more familiar content-bearing nouns. Section 4 proposes a semantics for certain *reason* and *way* sentences on the basis of this kinship. Section 5 provides the analytical payoff, demonstrating how these assumptions collectively predict definiteness in reason and manner descriptions.

¹Locations may also work in something like this way, but I'll set them aside for brevity.

2 Manners and reasons and the Paradoxical Cardinality Property

It's certainly not the case that reasons or manners are obligatorily singular or expressed with definite descriptions:

- (3) a. $\left. \begin{array}{l} \text{A} \\ \text{One} \\ \text{The (main)} \end{array} \right\}$ reason Floyd ran was fear.
- b. $\left. \begin{array}{l} \text{A} \\ \text{One} \\ \text{The (main)} \end{array} \right\}$ way in which Floyd ran was by taking huge strides.

All the forms in (3) give rise to the implicature that there were additional reasons apart from the one mentioned. This can be made an entailment as well:

- (4) a. He ran for two reasons: first, the swarm of bees, and second, the pack of hungry wolves.
- b. He ran in two ways: taking huge strides and bouncing from side to side.

Even quantification is possible:

- (5) a. He ran in every way I did.
- b. He ran in most ways I could think of.
- c. There is no way he can run (without injury).

Interestingly, *no way* is conventionalized to express emphatic negation. Omitting *without injury* from (5c) would tend to convey that he definitely can't run at all.

One striking property of ways and reasons is that they're hard to individuate. No matter what the facts of the matter are, it's hard to determine whether Floyd ran in one way or two or twelve. Likewise for reasons. That's true conceptually, but it's also clearly reflected linguistically. As far as I can see, (6a) and (6b) have the same truth conditions:

- (6) a. He ran for two reasons: first, the swarm of bees, and second, the pack of hungry wolves.
b. He ran for precisely one reason: the creatures pursuing him.

That's also the case for their manner counterparts:

- (7) a. He ran in two ways: taking huge strides and bouncing from side to side.
b. He ran in precisely one way: taking huge strides while bouncing from side to side.

There is an interesting side issue in (7) having to do with whether the manners are interpreted as describing a single event simultaneously or two distinct subevents. But this is an orthogonal feature of this particular example. Apparently, salsa dancing comes in two varieties, Puerto Rican and Cuban, which leads to sentences like these:

- (8) a. He danced in two ways: the Puerto Rican salsa and the Cuban salsa.
b. He danced in precisely one way: the salsa.

These can describe the same dancing event.

This difficulty of individuation is a special property of reasons and manners, and my suspicion is that it's the crucial one that explains their odd behavior with respect to definiteness. Fundamentally, there is no difference between two reasons and a single reason. More than that, they are cumulative, which is surprising for a notion expressed with a singular count noun. To lay this out a bit more fully, the extension of a singular count noun is not cumulative because the sum of any two objects in it is not also in it. Floyd and Clyde might both be in the extension of *linguist*, but their sum, the plural individual consisting of the two of them together, is not in the extension of *linguist*. But it is, of course, in the extension of *linguists*. And the extension of plural nouns *is* cumulative, because any two pluralities in the extension of *linguist* can be summed to make a new plurality that is in the extension of *linguists*. Mass nouns are similar: any two quantities in the extension of *water* can be summed to yield another quantity in the extension of *water*.

But that's not how manners and reasons work. One reason Floyd is running may be the angry bees. Another is the hungry wolves. Together, they are two reasons for him to run. It's therefore apparently enigmatic and surprising that together, they are also a single reason for him to run. That's true of manners as well. If Floyd is running taking huge strides and bouncing from side to side, these are two ways in which he's running, but they are also *the* way in which he's running.

For the sake of having a label, I'll call this the Paradoxical Cardinality Property of reasons and manners:

(9) **Paradoxical Cardinality Property**

The same event can be said to have a single reason, or arbitrarily many, and a single manner or way, or arbitrarily many.

The hypothesis we have arrived at links this to cumulativity:

(10) **Cumulativity Generalization**

The Paradoxical Cardinality Property arises for a singular noun N iff N is cumulative; that is, iff for any x and y in $\llbracket N \rrbracket$, the mereological sum of x and y is also in $\llbracket N \rrbracket$.

Any noun with this property would suspiciously resemble a plural or mass noun. But this shouldn't be alarming, and in some sense it's inevitable. Szabolcsi & Zwarts (1993) implicitly reach a similar conclusion for manners.

3 The wider world of paradoxical cardinality

As might be expected, *way*, *reason*, and their synonyms are not the only nouns with the Paradoxical Cardinality Property. There are various potential candidates for others, but a large class that presents itself is nouns that, it has been claimed, have propositional content, in the sense of Moulton (2009) and many others subsequently.

Idea is one such noun. First, the sense in which *idea* has propositional content is that, although it behaves compositionally like an individual, any idea is necessarily an idea *about* something. The usual properties we expect nouns to have are also properties of *idea*—it can occur in both singular and plural forms, with a wide range of quantifiers, and in a wide

range of nominal contexts. Treating it as having individuals in its extension therefore makes sense. But any given idea is an idea that something is the case. That's reflected in predicative sentences:

(11) Floyd's idea was that he shouldn't poke beehives anymore.

Not poking beehives can be said to be the propositional *content* of Floyd's idea.

Many other nouns denote properties of individuals with propositional content:

(12) Floyd's $\left. \begin{array}{l} \text{belief} \\ \text{thought} \\ \text{claim} \\ \text{assertion} \\ \text{allegation} \\ \text{accusation} \\ \text{suggestion} \end{array} \right\}$ was that he shouldn't poke beehives
anymore.

All of these have the Paradoxical Cardinality Property. There is a natural way of summing propositional content: with logical conjunction. The sum of the proposition that he shouldn't poke beehives anymore and that he shouldn't provoke wolves is a single proposition: that he shouldn't do one and he shouldn't do the other. If both of these propositions are beliefs of his, it's also necessarily the case that the conjunction of the two is a belief of his.² That's the case for all the content-bearing nouns in (12).

For the sake of explicitness, let's suppose that there is a sort of individual that is in the domain of a function, **content**, that maps individuals to their propositional content. It's therefore a function of type $\langle e, st \rangle$. An example:

²As a linguistic matter, in any case, this seems to be how we use nouns like *belief*. There is a philosophical debate about whether we all know the logical consequences of our beliefs. This is referred to as 'epistemic closure' (Luper 2020), a term since hijacked to describe isolated political media bubbles. I'm not sure to what extent the linguistic and philosophical issues here can be related.

- (13) a. Floyd’s idea was that he shouldn’t poke beehives.
 b. $\mathbf{content} \left(\iota x \left[\begin{array}{l} \mathbf{idea}(x) \wedge \\ \mathbf{Floyd}'s(x) \end{array} \right] \right) = \lambda w \left[\begin{array}{l} \text{Floyd shouldn't} \\ \text{poke beehives in } w \end{array} \right]$

Of course, it’s also sometimes necessary to determine the individual that has certain propositional content—the nominalized proposition, in the lingo. I’ll indicate that with the function **individual-counterpart**, which is of type $\langle st, e \rangle$:

- (14) a. The idea that Floyd shouldn’t poke beehives is wise.
 b. $\mathbf{wise} \left(\mathbf{individual-counterpart} \left(\lambda w \left[\begin{array}{l} \text{Floyd shouldn't} \\ \text{poke beehives in } w \end{array} \right] \right) \right)$

With this in place, a sum operation for content, \oplus_c , can be defined:

$$(15) \quad x \oplus_c y \stackrel{\text{def}}{=} \mathbf{individual-counterpart} \left(\lambda w \left[\begin{array}{l} \mathbf{content}(x)(w) \wedge \\ \mathbf{content}(y)(w) \end{array} \right] \right)$$

This says that the content sum of two individuals is the individual counterpart of the logical conjunction of their contents. For example, the content sum of the idea that Floyd shouldn’t poke beehives and the idea is that he shouldn’t antagonize snakes is the idea that he should do neither of these things. It’s slightly more elegant to state this in terms of sets as in (16), and this strategy will prove useful in a moment:

$$(16) \quad x \oplus_c y \stackrel{\text{def}}{=} \mathbf{individual-counterpart}(\mathbf{content}(x) \cap \mathbf{content}(y))$$

Of course, the content sum operation will not displace its more familiar cousins. We still need the classic Link (1983) sum operations, including the individual sum operation that combines singular (i.e., atomic) individuals to create plural individuals and the mass sum operation that combines bits of stuff in the extension of a mass noun to create larger agglomerations of stuff.

4 Manners, reasons, and contents

That’s all entirely independent from reasons and manners, so it remains to be seen whether these ideas will help. The notion of propositional content

instantly makes sense of sentences like (17a), which can be represented as in (17b), along the same lines as content copular sentences like (13):³

- (17) a. The reason for Floyd's flight is that he was pursued by a swarm of bees.
 b. $\mathbf{content}(\iota x[\mathbf{reason}(\mathbf{Floyd's-flight})(x)]) = \lambda w[\mathbf{Floyd\ was\ pursued\ by\ bees\ in\ }w]$

Manners can be treated analogously, with one small twist. The content of nouns like *manner* or *way* isn't a proposition. It's, well, a *manner*, a way of doing something. The standard way of construing manners since Davidson (1967) is to regard them as properties of events. The natural move, then, is to regard the content of nouns like *way* as properties of events as well. Thus:

- (18) a. The way Floyd fled is by taking huge strides.
 b. $\mathbf{content}(\iota x[\mathbf{way}(\mathbf{Floyd's-flight})(x)]) = \lambda e[\mathbf{Floyd\ took\ huge\ strides\ in\ }e]$

This is, in a sense, unsurprising. If events and worlds are both understood as species of situation in the Kratzer (1989) style, these two types of content are actually two sides of the same coin. The intersective semantics above for the content sum operation \oplus_c already makes possible summing content of this type.

From all this, it also follows that manners and reasons are cumulative, in the sense that the sum of two reasons is itself a reason and likewise for manners. That's because the relevant sum operation is content sum, which is structured to give rise to this through the intersective semantics of summing content.

5 The analytical payoff

But our aim was not to simply represent copular content sentences, but to explain the Paradoxical Cardinality Property and the unexpected definiteness of manner and reason DPs. These follow from the introduction of

³The constant **Floyd's-flight** has as its value the property of events of Floyd fleeing, or perhaps its individual counterpart.

the content sum operation. Maximal interpretations of plural definite descriptions arise from the fact that, with plurals, *the* picks out the maximal plural individual in the extension of the noun—strictly, its supremum. For the sake of explicitness, I'll write this with **sup**. But before finding the maximal individual in a predicate's extension, it has to be clear what sort of individuals are involved. If its extension consists of plural individuals, it's the maximal individual constructed with the usual individual sum operation, which I'll write **sup_i**. If its extension consists of mass individuals, it's the maximal individual constructed with the corresponding mass sum operation, **sup_m**. The innovation is in a third case. If its extension consists of content-bearing atomic individuals, it's the maximal individual constructed with the corresponding content sum operation introduced above, **sup_c**:

$$(19) \quad \llbracket the \rrbracket = \lambda P \begin{cases} \mathbf{sup}_i(P) & \text{if } P \text{ holds of plural individuals} \\ \mathbf{sup}_m(P) & \text{if } P \text{ holds of mass individuals} \\ \mathbf{sup}_c(P) & \text{if } P \text{ holds of atomic content-bearing individuals} \\ \iota(P) & \text{otherwise} \end{cases}$$

Thus when *the* combines with *reason* or *manner*, it picks out the individual with the largest content, the overall reason or manner. Naturally, contextual domain restrictions can constrain this in various contexts, as is the case for determiners in general.

Does this explain the Paradoxical Cardinality Property? I think so. In describing a single event, one can individuate its reasons and manners in arbitrary ways, just as one can divide a mass of water in arbitrary ways. For a particular event, *the reason* will pick out the reason with the largest content because $\llbracket reason \rrbracket$ holds of atomic content-bearing individuals, **sup_c**(*P*). But the same event can be said to have two reasons, or four or twenty. In these cases, *the reasons* will pick out the maximal plurality of reasons, **sup_i**(*P*). *Way* and *manner* work in precisely the same way. The paradoxical behavior of all these nouns arises because of the special character of content-bearing individuals.

References

- Davidson, Donald. 1967. 'The logical form of action sentences'. In Nicholas Rescher (ed.), *The Logic of Decision and Action*, pp. 81–95. University of Pittsburgh Press, Pittsburgh. Republished in Donald Davidson, 1980. *Essays on Actions and Events*. Oxford University Press, Oxford.
- Kratzer, Angelika. 1989. 'An investigation of the lumps of thought'. *Linguistics and Philosophy* **12**(5), 607–653.
- Link, Godehard. 1983. 'The logical analysis of plurals and mass terms: A lattice-theoretical approach'. In Rainer Bäuerle, Christoph Schwarze, & Arnim von Stechow (eds.), *Meaning, Use, and Interpretation of Language*. Walter de Gruyter, Berlin.
- Luper, Steven. 2020. 'Epistemic Closure'. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Summer 2020 edn.
- Moulton, Keir. 2009. *Natural Selection and the Syntax of Clausal Complementation*. Doctoral dissertation, University of Massachusetts Amherst.
- Szabolcsi, Anna & Frans Zwarts. 1993. 'Weak islands and an algebraic semantics for scope taking'. *Natural Language Semantics* **1**(3), 235–284.

Crosslinguistic variation in concessive scalar particles*

KIMIKO NAKANISHI
Meiji Gakuin University

1 Introduction

In Nakanishi and Rullmann (2009) and Rullmann and Nakanishi (2009), we examined semantic properties of what we call ‘concessive’ *at least* in English and the corresponding expressions in other languages. In (1), for example, English *at least* and Japanese *dake-demo* ‘(lit.) only-even’ has a ‘settle for less’ reading that although the speaker considers eating the ice cream to be less preferable than eating healthier food, he is content with it (because it is better than eating nothing).

- (1) a. Eat at least [the ice cream]_F.
b. [Aisu]_F-dake-demo tabe-nasai.
ice cream-DAKE-DEMO eat-IMP
(Rullmann & Nakanishi 2009, (28))

In our work, we left out the discussion of distributional differences. For instance, unlike *at least*, Japanese *dake-demo* is distributionally restricted: it is deviant in episodic sentences, as in (2), but licensed in downward-entailing (DE) contexts like in the restrictor of universal quantifiers (3) and the antecedent of conditionals (4), where it is generally glossed with *even* (Nakanishi 2006).¹

* My interests in concessive particles grew from working with Hotze. I admire his knowledge as a linguist and his warm heart as a person. I thank him for his insights and encouragement over the course of my research. Contact: kimiko@ltr.meijigakuin.ac.jp.

¹ Another context where NPIs appear is questions. Indeed, *dake-demo* is also felicitous in questions, as in (i), where it is glossed with *at least* (cf. Giannakidou 2007 on Greek *esto*). For lack of space, I do not discuss these examples any further.

- (i) [Aisu]_F-dake-demo tabe-ta-no?
ice cream-DAKE-DEMO eat-PAST-Q
‘Did you eat at least the ice cream?’ (Rullmann and Nakanishi 2009, (29))

- (2) *John-wa toi [iti]_F-dake-demo toi-ta.
 John-TOP question one-DAKE-DEMO answer-PAST
 ‘John answered even Question 1.’
- (3) [Toi [iti]_F-dake-demo toi-ta] dono hito-mo ukaru.
 question one-DAKE-DEMO answer-PAST everyone pass
 ‘Everyone who answers even Question 1 will pass.’
- (4) [Toi [iti]_F-dake-demo toi-ta]-ra ukaru.
 question one-DAKE-DEMO answer-PAST-if pass
 ‘If you answer even Question 1, you will pass.’

The distribution of *dake-demo* is not limited to the contexts where NPIs appear: *dake-demo* is felicitous in imperatives, as in (1b), and in a variety of modal contexts, as exemplified in (5) with the necessity modal. Just like in (1b), *dake-demo* in (5) carries a concessive interpretation.

- (5) John-wa toi [iti]_F-dake-demo toka-nebanaranai.
 John-TOP question one-DAKE-DEMO answer-must
 ‘John must answer at least Question 1.’

The distribution of *dake-demo* described here is the same as that of so-called ‘concessive scalar particles’ (CSPs) like Greek *esto* (Giannakidou 2007), Spanish *siquiera* (Alonso-Ovalle 2009, 2016) and *aunque sea* (Lahiri 2010), and Slovenian *magari* (Crnič 2011). CSPs are infelicitous in episodic sentences, but they can occur in various DE contexts, where they are glossed with *even*.² They can also appear in various modal environments, where they are glossed with *at least* and convey a concessive interpretation.

This short paper compares the distribution and interpretation of *dake-demo* with those of CSPs in other languages. Although *dake-demo* in principle patterns with other CSPs in its distribution, I demonstrate that there are some crucial differences. In particular, I show that *dake-demo* is truth-conditionally vacuous unlike CSPs in some other languages. Following Nakanishi (2006), I argue that *dake-demo* is composed of wide scope EVEN and narrow scope ONLY, and account for its distribution

² There is cross-linguistic variation as to under which DE contexts CSPs are licensed. For example, Spanish *siquiera* is licensed with clausemate negation (Alonso-Ovalle 2009, 2016), but *dake-demo* (Nakanishi 2006), Greek *esto* (Giannakidou 2007), and Slovenian *magari* are not (Crnič 2011).

and interpretation by introducing a new way of calculating the scalar presupposition of EVEN based on concessive conditional morphology.

2 Previous analyses

Regarding the question of why *dake-demo* is acceptable in DE contexts, but not in episodic sentences, I presented in Nakanishi (2006) an analysis that decomposes *dake-demo* into *dake* ‘only’ and *demo* ‘even’. This analysis is based on Guerzoni’s (2003) account of German *auch nur* ‘even’, which is an NPI that is licensed in DE contexts like (6), but not in episodic contexts.

- (6) Niemand hat auch nur [das Buch]_F gelesen.
 nobody has even the book read
 ‘Nobody even read the book.’

Guerzoni argues that *auch nur* consists of two focus-sensitive propositional operators associated with the same focus site, namely, *auch* ‘also’ (ALSO) in (7) and *nur* ‘only’ (ONLY) in (8). They take a contextually determined set of alternatives *C* as their first argument, and a proposition *p* as their second argument. They are truth-conditionally vacuous, but introduce presuppositions; ALSO triggers the existential presupposition (ExistP) that there is a proposition other than *p* that is true, while ONLY evokes two presuppositions: the exclusive presupposition (ExclusiveP) that there is no proposition other than *p* that is true, and the scalar presupposition (ScalarP) that *p* is the most likely among the alternatives in *C*.³

- (7) $[[\text{ALSO}]]^{g,c} = \lambda C. \lambda p. \lambda w: \exists q \in C [q \neq p \wedge q(w) = \text{True}]. p(w) = \text{True}$
- (8) $[[\text{ONLY}]]^{g,c} = \lambda C. \lambda p: \forall q \in C [q \neq p \rightarrow p >_{\text{likely}} q]. \lambda w: \neg \exists q \in C [q \neq p \wedge q(w) = \text{True}]. p(w) = \text{True}$

³ While ALSO has the same lexical meaning as regular additive particles, ONLY in (8) proposed for *auch nur* differs from regular exclusive particles. In particular, while regular exclusive particles are considered to make a truth-conditional contribution in terms of exclusivity (e.g., *Only Al came* is true iff there is no other person but Al who came), ONLY in (8) merely introduces presuppositions. This is based on the observation that *auch nur* is truth-conditionally vacuous (but see Schwarz 2005 for the claim that *auch nur* is an existential quantifier at the level of truth conditions).

The ExistP of ALSO and the ExclusiveP of ONLY are inconsistent, but the conflict can be resolved if there is an intervening DE operator. For instance, in (9), which is the LF of (6), *auch* presupposes that there is some book other than this book that nobody read, while *nur* presupposes that there is no book other than this book that everyone read (assuming that the trace is interpreted as a universal quantifier: see Heim 1983). These two presuppositions are consistent, correctly predicting (6) to be felicitous.

(9) ALSO C' [nobody₁ [ONLY C [t₁ read [[this book]_F]_F]]]

Extending Guerzoni's analysis, I argued in Nakanishi (2006) that *dake-demo* consists of *dake* 'only' (ONLY) that has the same lexical entry as *nur* in (8), and *demo* 'even' (EVEN) that has the same meaning as regular scalar particles, given in (10); EVEN makes no truth-conditional contributions, but it introduces the ScalarP that *p* is the least likely among the alternatives in *C*.

(10) $[[\text{EVEN}]]^{\text{sc}} = \lambda C. \lambda p: \forall q \in C [q \neq p \rightarrow q >_{\text{likely}} p]. \lambda w. p(w) = \text{True}$

With *dake-demo*, there is an inherent conflict between the ScalarP of EVEN and the ScalarP of ONLY. In the LF of (2) in (11), while EVEN presupposes that 'that John answered Q1' is the least likely, ONLY evokes the opposite ScalarP that the same proposition is the most likely. For the two ScalarPs to be consistent, there needs to be an intervening DE operator that reverses the scale of wide scope EVEN. With the LF: EVEN > DE > ONLY, the two ScalarPs yield the reading where Question 1 is taken to be the easiest, which is consistent with our intuition (see Section 4 for more discussion).

(11) EVEN C' [ONLY C [John answered Question [[one]_F]]]

As it stands, this compositional analysis can explain why *dake-demo* occurs in DE contexts, but not in episodic sentences. However, it cannot predict why *dake-demo* is licensed in various modal contexts. Presumably, in (5), the necessity modal intervenes between EVEN and ONLY, as in (12). Unlike DE operators, however, modals preserve the problematic entailments, and thus the two ScalarPs remain inconsistent. This predicts (5) to be infelicitous, contrary to the fact.

(12) EVEN C' [must [ONLY C [John answers Question [[one]_F]]]]

Crnič (2011) presents an analysis that can cover the full range of distribution of CSPs. He argues that CSPs like Slovenian *magari* consist of two focus-sensitive operators, EVEN and AT LEAST, that associate with the same focus site. While EVEN introduces the ScalarP in (10), AT LEAST is a weak existential quantifier, as in (13). In his analysis, (5) has the LF in (14).

$$(13) \quad \llbracket \text{AT LEAST} \rrbracket^{g,c} = \lambda C. \lambda p. \lambda w. \exists q \in C [p \geq_{\text{likely}} q \wedge (p(w) = \text{True} \vee q(w) = \text{True})]$$

(Crnič 2011:6)

$$(14) \quad \text{EVEN } C' \left[\text{must} \left[\text{AT LEAST } C \left[\text{John answers Question} \left[\left[\text{one} \right]_F \right]_F \right] \right] \right]$$

Suppose that there are three questions, Q1, Q2, and Q3. In this case, (5) has the strengthened free choice interpretation in (15a), where Q1 stands for the proposition that John answers Q1, etc. The alternatives for EVEN then are (15b), and since there are no entailment relations between the alternatives, the ScalarP of EVEN in (15c) is licit and captures the concessive meaning. The ScalarP in (15c) makes sense if Q1 is the easiest question among the three; it is more likely for the speaker to demand John to answer harder questions. The speaker settles for less by letting John to answer any of the questions, rather than requiring him to answer a harder question.

- (15) a. $\square(Q1 \vee Q2 \vee Q3) \wedge \diamond Q1 \wedge \diamond Q2 \wedge \diamond Q3$
- b. $\llbracket C' \rrbracket^{g,c} = \{ \square(Q1 \vee Q2 \vee Q3) \wedge \diamond Q1 \wedge \diamond Q2 \wedge \diamond Q3, \square(Q2 \vee Q3) \wedge \diamond Q2 \wedge \diamond Q3, \square Q3 \}$
- c. ‘that John must answer Q1 or Q2 or Q3 and John may answer Q1 and John may answer Q2 and John may answer Q3’ is the least likely among the alternatives in C’

Crnič’s analysis can also explain why CSPs are deviant in episodic contexts, but fine in DE contexts. Applied to (2), EVEN evokes the ScalarP that ‘that John answered Q1 or Q2 or Q3’ is the least likely, but this is contradictory. However, when EVEN scopes over a DE operator, the entailments get reversed, which makes the ScalarP plausible.

3 Cross-linguistic differences

One of the crucial differences between Nakanishi's (2006) analysis and Crnič's (2011) is the truth-conditional contributions of CSPs. The difference is apparent with universal modal examples such as (16a) with *magari*. Assuming that the relevant scale is <scanned photo, original photo>, the truth conditions under Crnič's analysis are (16b). Alonso-Ovalle (2016) claims that the corresponding example with Spanish *siquiera* has the same interpretation, supporting Crnič's claim that CSPs convey an existential meaning. In contrast, under my analysis as well as under any accounts that treat CSPs as being truth-conditionally vacuous (Giannakidou 2007; Alonso-Ovalle 2009; Lahiri 2010), (16a) is predicted to have the interpretation in (16c), which Crnič claims is too strong for Slovenian *magari* (and also for Spanish *siquiera*, as Alonso-Ovalle points out).

- (16) a. Za potni list mi mora Janez poslati magari
 for passport me must John send magari
 [poskenirano]_F sliko.
 scanned photo
 'To get a passport, John must send me at least a scanned photo.'
 (Crnič 2011:5)
- b. $\square(\text{scanned.photo} \vee \text{original.photo}) \wedge \diamond\text{scanned.photo} \wedge \diamond\text{original.photo}$
- c. $\square\text{scanned.photo}$

Unlike *magari* and *siquiera*, however, *dake-demo* is indeed truth-conditionally vacuous; unlike (16a), (17) with *dake-demo* is infelicitous. This can be explained if the assertion of (17) is (16c). The requirement in (16c) is pragmatically odd; it is hard to imagine a situation where a scanned photo, but not an original one, is needed for a passport application.

- (17) #Pasupooto sinsei-notame John-wa [sukyansita]_F
 passport application-for John-TOP scanned
 syasin-dake-demo okura-nebanaranai
 picture-DAKE-DEMO send-must
 'For a passport application, John must send at least a scanned photo.'

Similarly, in (5) with *dake-demo*, John is required to answer Q1; he may possibly answer other questions in addition, but the asserted requirement is for him to answer Q1. In contrast, in the corresponding examples with *magari* or *siquiera*, John may answer any of the questions, as in (15a). Likewise, in the imperative (1b) with *dake-demo*, the addressee is required to eat ice cream.

The same claim holds for *dake-demo* in DE contexts. In (3), to pass, it is necessary for everyone to answer Q1; it is not certain whether they pass by answering Q2 or Q3 without answering Q1, even if Q2 and Q3 are harder than Q1. Similarly, in (4), the addressee needs to answer Q1 in order to pass. There is no guarantee of his passing by answering harder questions without answering Q1.⁴

In sum, there is a cross-linguistic difference as to the truth-conditional contributions of CSPs. Unlike *magari* or *siquiera*, *dake-demo* makes no contributions. Thus, any analysis that allocates a CSP a weak existential meaning (such as Crnić 2011 or Alonso-Ovalle 2016) is not suitable for *dake-demo*. In contrast, the correct truth conditions can be derived from Nakanishi's (2006) compositional analysis presented in the previous section. However, as pointed out above, this analysis fails to account for why *dake-demo* is licensed in modal contexts. In the following, I suggest a way of saving the analysis by adopting a new way of calculating the ScalarP of EVEN.

4 Proposal

CSPs are licensed with the necessity modal, as shown in (18a) with Spanish *aunque sea* and (18b) with Japanese *dake-demo*. Assuming that the LF of (18) is (19), the ScalarP of EVEN says that 'that you must go to the doctor once a month' is the least likely. As Lahiri (2010) points out, this presupposition is not correct: the target proposition is entailed by 'that you must go to the doctor n times a month' (n>1). This is because *must* preserves the problematic entailment, as pointed out above. Then

⁴ The observation here shows that *dake-demo* lacks what Schwarz (2005) calls 'characteristic implications', observed with German *auch nur*. For instance, (ii) implies that the speaker loses the bet if Hans read the second or third volume.

(ii) Wenn Hans auch nur den [ersten]_F Band gelesen hat, dann verliere
 if Hans even the first volume read has then lose
 ich die Wette.
 I the bet

'If Hans has even read the first volume, I lose the bet.' (Schwarz 2005:151)

the target proposition is taken to be the most likely, which wrongly predicts (18a,b) to be unacceptable.

- (18) a. Tiene usted que ir al médico aunque sea
 have.to you go to the doctor “even”
 [una]_F vez al mes.
 once a month
 ‘You have to go to the doctor at least once a month.’
 (Lahiri 2010:20)

- b. Isya-ni tuki [ik-kai]_F-dake-demo
 doctor-DAT month one-time-DAKE-DEMO
 ika- nebanaranai
 go-must
 ‘You have to go to the doctor at least once a month.’

- (19) EVEN C’ [must [ONLY C [you go to the doctor [[once]_F]_F a month]]]

Faced with this problem, Lahiri (2010) claims that *aunque sea* in the modal contexts is a narrow scope operator interpreted in the scope of the modal, but its ScalarP is calculated based on a conditional statement whose antecedent is the proposition that *aunque sea* applies to and whose consequent is some contextually salient goal. In (18a), for instance, there is a contextually salient goal such as that you will stay healthy. *Aunque sea* combines with the proposition ‘that you go to the doctor once a month’, and evokes the ScalarP that ‘that if you go to the doctor once a month, you will stay healthy’ is the least likely. This is plausible; the likelihood of leading a healthy life increases as the number of your visits to the doctor increases. However, Lahiri shows that this analysis faces problems when applied to *aunque sea* in DE contexts. He thus concludes that two different analyses are required. Below I search for a way of maintaining a unified analysis of *dake-demo*.

Lahiri’s (2010) analysis in modal contexts is motivated by the fact that *aunque sea* has concessive conditional morphology: *aunque* is one of the particles used to form *even if* conditionals, and *sea* is the third person singular present subjunctive of *ser* ‘be’. For instance, the concessive conditional in (20) is formed with *aunque*, and *sea* serves as the main verb of the antecedent clause, agreeing with the *pro* subject. As Lahiri notes, the literal translation of *aunque sea* is ‘even if (it) be.SUBJ’.

- (20) Aunque sea podrida, tienes que recoger la
 even.if be.3SG.SUBJ rotten.SG have.2SG to pick the
 manzana.
 apple

‘You have to pick the apple, even if it is rotten.’

(Lahiri 2010:14)

Japanese *demo* in

(21) is generally considered to be a focus particle corresponding to *even*, but it may be analyzed as having a clausal structure consisting of the copula *de* ‘be’ and the particle *mo* ‘also, even’ (Hiraiwa & Nakanishi 2021; Nakanishi 2021; Oda 2021). More specifically,

(21) may be viewed as a concessive conditional whose antecedent has the main verb *de* ‘be’ with the *pro* subject. That is,

(21) is paraphrasable to ‘Even if (it) be a child, (he) will pass.’

- (21) Kodomo-demo ukaru.
 child-DEMO pass
 ‘Even a child will pass.’

Note here that the conditional contains two phonologically silent pronouns. The one in the main clause is posited in place of the expression to which *demo* attaches, and thus its interpretation corresponds to that expression (in

(21), *kodomo* ‘child’). Regarding the one in the antecedent, it denotes an individual concept (Romero 2005), assuming that the copula sentence is specificational (in the sense of Higgins 1973). Roughly, it is interpreted as an individual that is salient in the context (cf. Oda 2021), that is, whatever individual that is freely picked up by the appropriate context, just like Lahiri’s goal (in

(21), the person who takes the relevant exam, etc.).

I continue to assume that *demo* is a scalar particle that is truth-conditionally vacuous and introduces a ScalarP just like *even*. Taking the spirit of Lahiri’s (2010) analysis, I propose that its presupposition is calculated based on a conditional statement. However, the content of the conditional statement is different from Lahiri’s; in my analysis, the

antecedent is a specificational copula sentence, and the consequent is the proposition that EVEN combines with. In

(21), for example, the ScalarP is that ‘if (it) be a child, (he) will pass’ is the least likely, where the subject in the antecedent is interpreted as something like the person who takes the exam. This ScalarP is plausible when other people are more likely to pass, which is consistent with our intuition.

Regarding *dake-demo*, I maintain Nakanishi’s (2006) analysis that it consists of EVEN and ONLY, both of which introduce presuppositions without making contributions to the truth conditions. The only difference is how the ScalarP of EVEN is calculated. The LF of (18b) is thus the one in (19). The sentence asserts that the addressee is required to go to the doctor once a month, and it has the presuppositions of ONLY and EVEN. ONLY introduces the ScalarP that visiting the doctor once a month is the most likely, i.e., it is easier than visiting multiple times. The ScalarP of EVEN is calculated based on a conditional statement; it says that ‘if the number of your possible visits to doctor is once a month, you must go to the doctor that many times’ is the least likely. This ScalarP is satisfied; the conditional statement that EVEN combines with entails the alternatives of the form ‘if the number of your possible visits to the doctor is n times a month, you must go to the doctor n times’ ($n > 1$), which means that the former is the least likely. Furthermore, this ScalarP together with the assertion successfully captures the concessive interpretation; the speaker is more likely to require the addressee to go to the doctor as often as possible (which is harder than going just once), but he settles for less by demanding the addressee to go just once (as stated in the assertion).⁵

Let us now see whether the analysis extends to *dake-demo* in DE contexts. Take (3), for example, whose LF is provided in (22). The ScalarP of ONLY says that ‘that x answers Q1’ is the most likely (where x is universally quantified, following Heim 1983), which suggests that Q1 is easier than other questions. The ScalarP of EVEN applied to a conditional statement says that ‘if the question on the test is Q1, everyone who answers it will pass’ is the least likely. This is sensible: the likelihood that everyone who answers Q1, which is the easiest (as the

⁵ Alonso-Ovalle (2009, 2016) presents an example with Spanish *siquiera* corresponding to (18), reporting that the sentence is acceptable even when the speaker is less likely to require the addressee to go more often than once a month. However, (18b) with *dake-demo* is infelicitous in such a context. In particular, in (18b), there is a strong expectation on the speaker’s part that the addressee goes to the doctor as often as possible.

ScalarP of ONLY suggests), will pass is less than the likelihood that everyone who answers harder questions will pass.

- (22) EVEN C' [everyone₁ [ONLY C [t₁ answers Question [[one]_F]]] will pass]

We can also explain why (3) is infelicitous when *ukaru* 'pass' is replaced by *otiru* 'fail' (see Lahiri 2010 for this pattern). According to the ScalarP of EVEN, 'if the question on the test is Q1, everyone who answers it will fail' is the least likely, but this is only licit in a pragmatically odd context where the likelihood of failing increases by answering harder questions.

Let us now examine *dake-demo* in episodic contexts. When *demo* is used in episodic sentences like (23), it is interpreted as a concessive without a hypothetical meaning (Tomura 1988, among others); (23) is paraphrasable as 'Although (it) was a child, (he) passed', where two pronouns are interpreted in the same way as

- (21) (e.g., *it* is something like the person who took the exam).⁶

- (23) Kodomo-demo ukat-ta.
 child-DEMO pass-PAST
 'Even a child passed.'

Based on these morphological data, I propose that in episodic contexts like (23), the ScalarP of *demo* is calculated based on a concessive statement. In (23), the ScalarP is that 'although (it) was a child, (he) passed' is the least likely. This is plausible and also consistent with our intuition.

The ScalarP applied to a concessive statement can account for why *dake-demo* is infelicitous in episodic sentences, as in (2), whose LF is given in (11). The ScalarP of EVEN is that 'although the question on the test was Q1, John answered it' is the least likely. This is sensible in the context where Q1 is taken to be the hardest question. However, this is inconsistent with the ScalarP of ONLY that 'that John answered Q1' is the most likely, which suggests that Q1 is the easiest. Indeed, just like (23), (2) without *dake* is acceptable when Q1 is considered to be the hardest question.

⁶ Similarly, while Spanish *aunque* with a subjunctive clause expresses the sense of 'even if', as in (20), it expresses the sense of 'even though' when used with an indicative clause (Lahiri 2010, see also Haspelmath & König 1998).

Finally, the current analysis can further account for the observation that when *dake-demo* is used with an actual event in the past, as in (2) and (24) with *toi-ta* ‘answer-PAST’, the infelicitous sentences become perfectly acceptable when the agent’s emotion (such as a desire or a regret) is explicitly expressed, as exemplified in (24) (Yamanishi 2015). In the latter case, EVEN evokes the ScalarP that ‘although the question on the test was Q1, I wanted to answer it’ is the least likely. This is plausible; it is more likely for the speaker to want to solve harder questions, assuming that Q1 is the easiest question (as the ScalarP of ONLY suggests).

- (24) Toi [iti]_F-dake-demo {*toi-ta / tok-itakat-ta}.
question one-DAKE-DEMO answer-PAST answer-want-PAST
‘I {answered/wanted to answer} at least Question 1.’

5 Concluding remarks

This paper examined the distribution and interpretation of *dake-demo*. Following Nakanishi (2006), I argued that *dake-demo* consists of wide scope EVEN and narrow scope ONLY, which introduce presuppositions without contributing to the truth conditions. I further proposed a novel way of calculating the ScalarP of EVEN, applied to a conditional or a concessive statement, and by so doing accounted for the distribution of *dake-demo* as well as concessive interpretations in modal contexts.

The current work sheds light on two cross-linguistic variations in CSPs, namely, whether CSPs convey an existential meaning, and whether CSPs have concessive conditional morphology. I consider the second of the two to be especially important as little investigation has been done from a cross-linguistic perspective. Further investigation is required to determine the relevance of such morphology to the distribution and interpretation of CSPs.

References

- Alonso-Ovalle, Luis. 2009. *Even* and biased questions: The case of Spanish *siquiera*. *SALT* 19:1–18.
- Alonso-Ovalle, Luis. 2016. Are all concessive scalar particles the same? Probing into Spanish *siquiera*. *SALT* 26:185–204.
- Crnič, Luka. 2011. On the meaning and distribution of concessive scalar particles. *NELS* 41:143–156.

- Giannakidou, Anastasia. 2007. The landscape of EVEN. *Natural Language & Linguistic Theory* 25:39–81.
- Guerzoni, Elena. 2003. *Why even ask? On the pragmatics of questions and the semantics of answers*. Ph.D. dissertation, MIT.
- Haspelmath, Martin & Ekkehard König. 1998. Concessive conditionals in the languages of Europe. In Johan van der Auwera (ed.), *Adverbial Constructions in the Languages of Europe*. Berlin, Germany: Mouton, 563–640.
- Heim, Irene. 1983. On the projection problem for presuppositions. *WCCFL* 2:114–125.
- Higgins, Francis R. 1973. *The Pseudocleft Construction in English*. Ph.D. dissertation, MIT.
- Hiraiwa, Ken & Kimiko Nakanishi. 2021. Japanese free choice and existential indeterminates as hidden clauses. *WAFJL* 15:37–44.
- Lahiri, Utpal. 2010. Some *evens* are *even (if) ... only*: The concessive “*even*” in Spanish. Manuscript.
- Nakanishi, Kimiko. 2006. *Even, only*, and negative polarity in Japanese. *SALT* 16:138–155.
- Nakanishi, Kimiko. 2021. A clausal analysis of free choice *demo* in Japanese. *Proceedings of the Linguistic Society of America (PLSA)* 6(1).
- Nakanishi, Kimiko & Hotze Rullmann. 2009. Epistemic and concessive interpretations of *at least*. Presented at *The Meeting of the Canadian Linguistics Association*, Carleton University.
- Oda, Hiromune. 2021. Japanese free choice items as unconditionals. *NALS* 29:281–338.
- Romero, Maribel. 2005. Concealed questions and specificational subjects. *Linguistics and Philosophy* 28:687–737.
- Rullmann, Hotze & Kimiko Nakanishi. 2009. More about *at least*. Presented at *MOSAIC (Meeting Of Semanticists Active In Canada)*, University of Ottawa.
- Schwarz, Bernhard. 2005. Scalar additive particles in negative contexts. *NALS* 13:125–68.

- Tomura, Kayo. 1988. Nihongo-niokeru hutatu-no taipu-no joojobun [Two types of concessives in Japanese]. *Studies in Language and Literature* 15:123–133.
- Yamanishi, Masako. 2015. ‘Dakedemo’ koo [On *dakedemo*]. *Mejiro Journal of Humanities* 11:249–260.

Sentential negation in Brazilian Portuguese*

MICHEL NAVARRO

University of British Columbia

1 Introduction

This paper offers a syntactic-semantic analysis of sentential negation in (standard) Brazilian Portuguese (BP). I argue that the negation marker *não* ‘not’, commonly assumed to be the semantic negation in BP (Schwenter 2005; Sousa 2007, 2011; Lopes & Rocha 2017), is in fact a negative element morpho-syntactically marked for negation, but semantically vacuous. Based on a range of empirical evidence, I show that the negative element *não* ‘not’, as the head of a phrase projected from its merge with VP, forms a complex head with V^0 and I^0 . Given this analysis, I demonstrate that if *não* were the semantic negation, it could only be a predicate negation. One immediate unwanted consequence of this is that the scope interactions between negation and universal quantification in BP would be left unaccounted for. Drawing on Zeijlstra’s theory of Negative Concord as a syntactic agree relation between a single interpretable feature [iNEG] and one or multiple uninterpretable features [uNEG], I propose that *não* is the head of a Polarity Phrase (PolP) generated by its merge with VP. In this approach, *não* hosts a [uNeg] feature and as such is licensed by the insertion of a covert NEG operator above IP. The advantage of this analysis is that it accounts for the scope ambiguity of sentences in BP with negation and universal quantification, as opposed to an analysis that treats the negative word *não* as the semantic negation.

The paper proceeds as follows. In Section 2, I offer a range of syntactic evidence that the negative element *não* in BP is the head of a maximal projection that selects a VP as complement. It is also demonstrated that *não* forms a complex head with V^0 and I^0 . In Section 3, I show that by treating *não* as a semantic negation head of a NegP, we would be committed to interpreting it as denoting a predicate negation. As a result, sentences where it co-occurs with the universal quantifier

* Firstly, I’d like to thank Prof. Hotze Rullmann for his impactful contributions to the field of linguistics and for being a model of a researcher, professor, and teacher. I’d like also to express my immense gratitude to him for accepting to be my Ph.D. supervisor. His constant guidance, insightful thoughts, and valuable comments have had an immeasurable impact on the progression of my research.

todo will always be represented as having truth conditions in which the negation has narrow scope. This contradicts the speakers' judgements, which allow for wide and narrow scope interpretations. In Section 4, I propose an analysis of *não* as an element that is not semantically marked for negation and occupies the head of a Polarity Phrase (PolP). In this configuration, it establishes a formal agreement relation with an unpronounced semantically active operator that c-commands it, and which, à la Rizzi (1996), is in the Spec of a NegP above IP. This approach correctly captures the scope ambiguity left unexplained in Section 3. Finally, Section 5 is the conclusion.

2 The distribution of *não*

In standard BP, the negation marker *não* always occurs pre-verbally, as shown in (1) and (2):


- (1) a. Alberto **não** ama Maria.
 Alberto not love.PRES.3SG Maria
 'Alberto doesn't love Maria.'
- b. *Alberto ama **não** Maria.
- (2) a. Quem Alberto **não** ama?
 who Alberto not love.PRES.3SG
 'Who doesn't Alberto love?'
- b. *Quem Alberto ama **não**?

Furthermore, *não* immediately precedes the verb.¹ So much so that (aspectual) adverbs, such as *frequentemente* 'often', although having a sentential distribution relatively unconstrained in BP, as shown in (3), are blocked from intervening between *não* and the verb, as illustrated by the ungrammaticality of (4):

¹ In this paper, I don't consider a variety of BP, spoken in Northeastern Brazil, in which *não* can be post-verbal. Furthermore, I will not analyse emphatic uses of *não*, i.e., clauses in which there are two co-occurrences of *não*: one pre-verbal and the other clause-final, the latter being the emphatic negation. I believe these cases can be incorporated into my analysis. However, for reasons of space, they are not investigated here. For pragmatic analyses of them see Schwenker 2005 and Sousa 2011.

- (3) a. *Frequentemente* Maria **não** visita Sandra.
 often Maria not visit.PRES.3SG Sandra
 ‘Maria doesn’t often visit Sandra.’
- b. Maria *frequentemente* **não** visita Sandra.
 Maria often not visit.PRES.3SG Sandra
 ‘Maria doesn’t often visit Sandra.’
- c. Maria **não** visita *frequentemente* Sandra.
 Maria not visit.PRES.3SG often Sandra
 ‘Maria doesn’t often visit Sandra.’
- d. Maria **não** visita Sandra *frequentemente*.
 Maria not visit.PRES.3SG Sandra often
 ‘Maria doesn’t often visit Sandra.’
- (4) *Maria **não** *frequentemente* **visita** Sandra.
 Maria not often visit.PRES.3SG Sandra
 ‘Maria doesn’t often visit Sandra.’

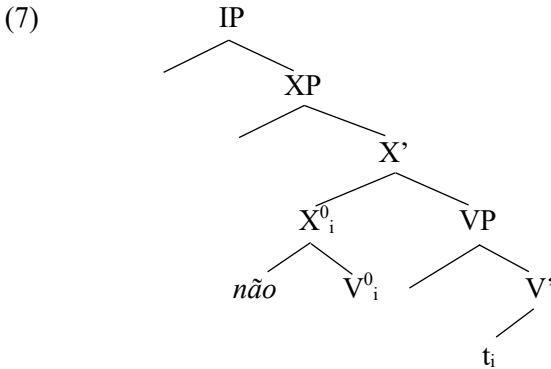
On the assumption that adverbs (in BP) are base generated in a position demarcating the left boundary of the VP (Pollock 1989), as in (5), I suggest that (6a), the positive counterpart of (4), is evidence of V^0 -to- I^0 raising in BP, as in (6b).²

- (5) [IP NP I [VP (**Adv**) V...]]
- (6) a. Maria visita *frequentemente* Sandra.
 Maria visit.PRES.3SG often Sandra
 ‘Maria often visits Sandra.’
- b. [IP visita_i [VP (**frequentemente**) ...t_i...]]


Given (5) and the account of (6a) just outlined, I propose that (4) is ruled out because the presence of *não* as the head of an XP projected from the merge of *não* with a VP triggers the raising of V^0 to X^0 , forming a complex head with it and blocking the intervention of an adverb between

² For a classical overview of the head movement debate, see Roberts 2003.

them. Thus, under this view, the structure *não*+V⁰ in BP is represented as in (7):



As a result, whenever there is an adverb adjoined to V', the head *não*+V⁰ is pronounced above it. In turn, the tense marker hosted by I⁰ triggers V⁰ or *não*+V⁰ movement to I⁰. Such an analysis assigns, hence, the syntax (8b) to *não*+V⁰ raising in sentences of the type (8a):

- (8) a. Maria **não visita** *frequentemente* Sandra.
 Maria not visit.PRES.3SG often Sandra
 ‘Maria doesn’t often visit Sandra.’

- b. [IP ***não*+V⁰**_i [XP ... t_i ... [VP (Adv) ... t_i ...]]]

Another piece of evidence indicating that *não* forms a complex head with V⁰ comes from negative questions where I⁰+*não*+V⁰ precedes the subject. Consider sentence (9):

- (9) O que **não viu** Pedro?
 D what not see.PAST.3SG Pedro
 ‘What didn’t Pedro see?’ *wh*-word_{object} I⁰+***não*+V⁰** S

By contrast, V⁰ cannot raise and leave behind *não*, as in (10):

- (10) *O que **viu** **não** Pedro?
 D what see.PAST.3SG not Pedro
 ‘What didn’t Pedro see?’ **wh*-word_{object} I⁰+V⁰ ***não*** S

Assuming that the subject is in the Spec of IP, the well-formedness of (9) and the ungrammaticality of (10) strongly support the hypothesis that in (9) $I^0+n\tilde{a}o+V^0$ can raise to the head of a YP above IP. In due time (Section 4), after spelling out my analysis of sentential negation as an unpronounced operator in the Spec of a NegP, I will argue that the landing site Y^0 of $I^0+n\tilde{a}o+V^0$ in sentences where *não* and verbs precede the subject is the head of NegP. But before that, let us look at the consequence for scope ambiguity if we assume that *não* is the semantic negation in BP.

3 Scope ambiguity and *não*

In a negative sentence of the type (11a), which exhibits the linear order $S\ n\tilde{a}o+V$, the scope interaction between the generalized quantifier (GQ) *todo mundo* ‘everybody’ and negation gives rise to an ambiguity between two readings. In one reading, the universally quantified DP has scope over the negation, as paraphrased in (11b), abstracting away from tense. Another reading is one in which the negation scopes over the GQ, as captured by (11c).

- (11) a. **Todo mundo não** chegou.
 everybody not arrive.PAST.3SG
 ‘Everybody didn’t arrive/hasn’t arrived.’
- b. Reading 1: $\forall x[\text{person}_C(x) \rightarrow \neg \text{arrive}(x)]$ ($\forall > \neg$)
 c. Reading 2: $\neg \forall x[\text{person}_C(x) \rightarrow \text{arrive}(x)]$ ($\neg > \forall$)

Reading 1 is true in a scenario where nobody arrived, whereas reading 2 can describe a situation in which some people arrived, and some didn’t. In (11b) and (11c) the restrictor of the universal quantifier is a contextually salient set of people. I assume that C is a variable introduced by the quantifier that ranges over a salient subset provided by each context of utterance (von Stechow 1994).

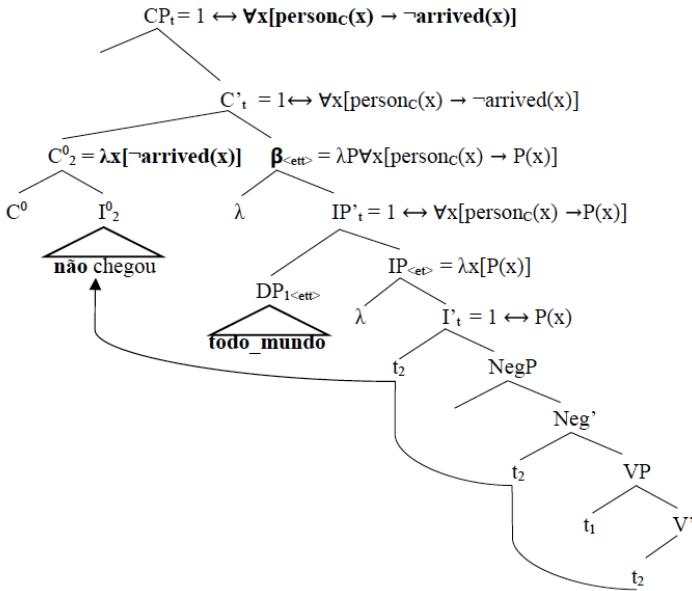
Importantly, the raising of $I^0+n\tilde{a}o+V^0$, resulting in an inverted order, as in (12), doesn’t seem to rule out the scope ambiguity attested above.

- (12) **Não** chegou **todo mundo**.
 not arrive.PAST.3SG everybody.
 ‘Didn’t arrive/hasn’t arrived everybody.’ ($\neg > \forall / \forall > \neg$)

Recall that we have demonstrated in Section 2 that *não* always forms a complex head with V^0 and I^0 . As such, if one assumes that *não* is the semantic negation, she must provide a compositional semantics for (11a) and (12) that captures their available readings. However, such an analysis fails the task from the start. To show this, let us temporarily assume that *não* is the head of a NegP between VP and IP. On top of that, I posit, for the sake of the argument, that $I^0+não+V^0$ covertly raises to C^0 in (11a), whereas in (12) it raises overtly. Lastly, let us assume that the GQ *todo mundo*, by Quantifier Raising (QR), adjoins to IP. Such a configuration is exhibited in (13), which is cast in conventions of Heim and Kratzer (1998).³

- (13) a. $[_{CP} [_{C'} [C^0 \text{ não} \text{ chegou}_2 [_{IP'} \text{ todo_mundo}_1 [_{IP} t_1 t_2 [_{NegP} t_1 t_2 [_{VP} t_1 t_2]]]]]]]]]$

b.



In (13b), the arrow indicates the stages of head movement of V^0 to Neg^0 , then to I^0 and finally to C^0 . The indexed t_2 indicates the traces left behind by the movements, which are assumed to be of the same type as the raised head(s). As for t_1 , it is the trace of type e left by the QR of *todo mundo*.

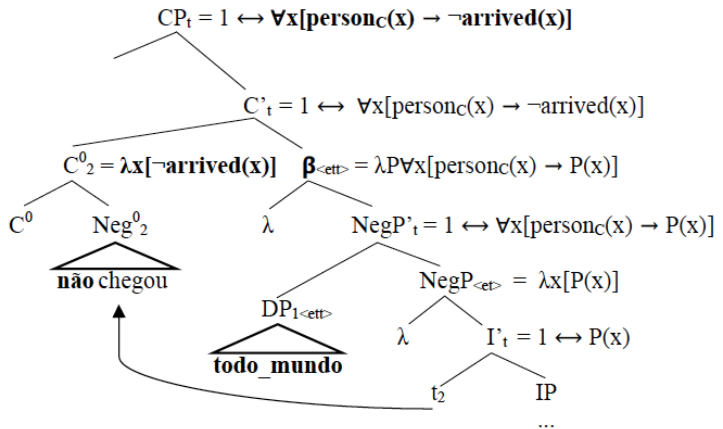
³ For a detailed elaboration of the QR theory, see May 1977, 1985.

By following the (partial) derivation illustrated above, it is easy to see that the truth conditions generated are ones in which the GQ has scope over the negation. This is so because *não*, due to its syntactic status as part of the complex head $C^0+I^0+não+V^0$, can only be a predicate negation, i.e., it can only take predicates as arguments. One crucial element here is that the movement of the complex head does not affect the interpretation. This is so because, by leaving a higher-order variable as their traces, predicative complex heads containing V^0 undergo semantic reconstruction. In other words, although the complex head $C^0+I^0+não+V^0$ c-commands the DP *todo mundo*, the former inevitably is interpreted as the semantic argument of the universal quantifier.⁴ As a consequence, the only reading is one with *não* having narrow scope with respect to *todo mundo*.

The same result holds if one assumes that *não* merges with IP. Let's posit that in this case the complex head $não+I^0+V^0$ undergoes head movement to C^0 and by QR the GQ *todo mundo* adjoins to NegP, as laid out in (14).

- (14) a. $[_{CP} [C' [C^0 \text{ não } \text{chegou}_2 [_{NegP'} \text{ todo_mundo}_1 [_{NegP} [IP \ t_1 \ t_2 [_{VP} \ t_1 \ t_2]]]]]]]$

b.



⁴ The analysis of head movement as a PF operation and as such semantically null (see Chomsky 1995, 2001; Schoorlemmer & Temmerman 2012; LaCara 2016) would lead to the same result as the one offered above, since in both accounts head movement does not affect the semantic interpretation.

As can be easily seen, even though the complex head $C^0+n\tilde{a}o+I^0+V^0$ again c-commands *todo mundo*, the outcome is the same truth conditions as the ones exhibited by (13a–b), with the negation under the scope of the universal quantifier. No matter whether the complement of *n\tilde{a}o* is a VP or IP, by the very fact that it forms a complex head with I^0 and V^0 , *n\tilde{a}o* cannot help but be a verbal predicate negation, hence forcing a narrow scope interpretation.

To summarize, due to its syntax, *n\tilde{a}o* is interpreted as negating a verbal predicate. That is, *n\tilde{a}o* is a function that takes a predicate as argument and returns a predicate as value. As a result, even in a configuration where the complex head c-commands the universally quantified DP, semantic reconstruction will give rise to the narrow scope reading of *n\tilde{a}o*. Thus, if the premise that *n\tilde{a}o* is part of a complex head is true, one available reading of (11a) and (12) is left unaccounted for by any compositional theory that treats *n\tilde{a}o* as hosting the semantic negation. But there is a way out of this stalemate, and it does not amount to giving up the analysis in Section 2. A solution will be offered in the next section.

4 The analysis

4.1 Theoretical background

In this section, in line with Ladusaw’s (1992) analysis of negation in English, I propose that sentential negation in BP is accomplished via an unpronounced NEG operator and that *n\tilde{a}o* is a negative element morpho-syntactically marked for negation, but semantically vacuous. My analysis also builds on Zeijlstra’s theory of Negative Concord as an instance of syntactic agreement (2004, 2008, 2012). I claim that *n\tilde{a}o*, as a semantically non-negative word, carries an uninterpretable feature [uNEG] that is checked in an upward agree relation it establishes with an abstract single NEG operator (above IP) that carries the interpretable feature [iNEG] and c-commands it. That is, *n\tilde{a}o* is licensed by the NEG operator.

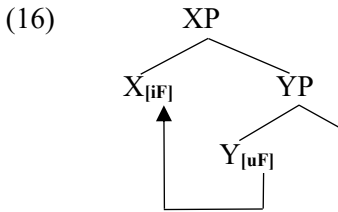
Two theoretical assumptions are crucial to oil the wheels of my analysis. Firstly, I adopt Zeijlstra’s Upward Agree condition below, which reverses the canonical direction of agreement relations:

(15) **Upward Agree** (Zeijlstra 2004, 2008, 2012)

α can agree with β iff:

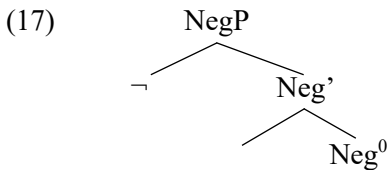
- a. α (Probe) carries at least one uninterpretable feature [uF] and β (Goal) carries a matching interpretable feature [iF].
- b. β c-commands α .
- c. β is the closest goal to α .

Upward Agree states that agreement — defined as a relation between an element that carries an interpretable formal feature and one or more element(s) that have uninterpretable counterparts of this same feature (Chomsky 1995, 2001) — is established via a c-command relation between a goal and a probe constituent. The goal hosts the interpretable feature [iF] and c-commands the probe, i.e., the constituent carrying the uninterpretable feature [uF] that must be checked for the derivation to converge. Formal agreement, according to (15), presents the following configuration:



Crucially, on this approach, it is not mandatory that the constituents with the uninterpretable features move in order for agreement to occur. Once Y has the uninterpretable counterpart of the interpretable feature that X carries, by the very fact that X c-commands Y, [uF] is immediately checked. Furthermore, Upward Agree allows for multiple uninterpretable features in the same clause to be checked against one interpretable feature, as is the case with Negative Concord (NC) constructions.

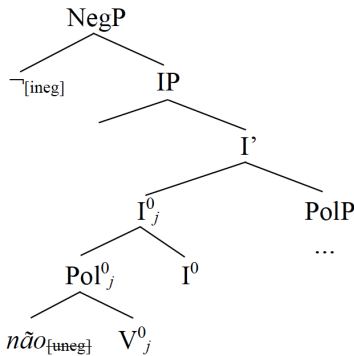
Secondly, I assume Rizzi’s Negative Criterion, which posits that (i) a semantic operator such as NEG “fills an A-bar specifier position” (Rizzi 1996:74) in a Spec-head configuration, and (ii) “negative sentences involve an independent clausal projection, the Negative Phrase” (Rizzi 1996:74). Thus, the structure proposed is as follows:



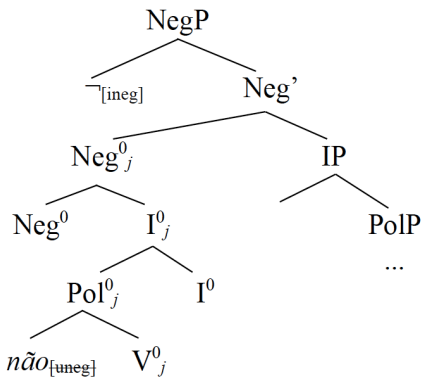
4.2 NEG in BP

Now, with the apparatus laid out above, which brings together the Upward Agree condition and the Negation Criterion, we can consider sentences (11a) and (12) again and look at how their common scope ambiguity can be accounted for. I argue that the sentential NEG operator hosting the [iNEG] feature in BP is in the Spec of NegP, just above IP and that *não*, analysed as a semantically vacuous negative marker that carries an uninterpretable feature [uNeg], is the head of a Polarity Phrase (PolP) right above VP. Therefore, having in mind the syntax of *não* offered in Section 2, I assign two configurations to negative sentences with *não* in BP: one in which $\text{Pol}^0 + \text{V}^0$ lands in I^0 , as is the case in sentence (11a), and one in which $\text{V}^0 + \text{Pol}^0 + \text{I}^0$ raises to Neg^0 , as in (12a). The former is displayed in (18a) and the latter in (18b):

(18) a.



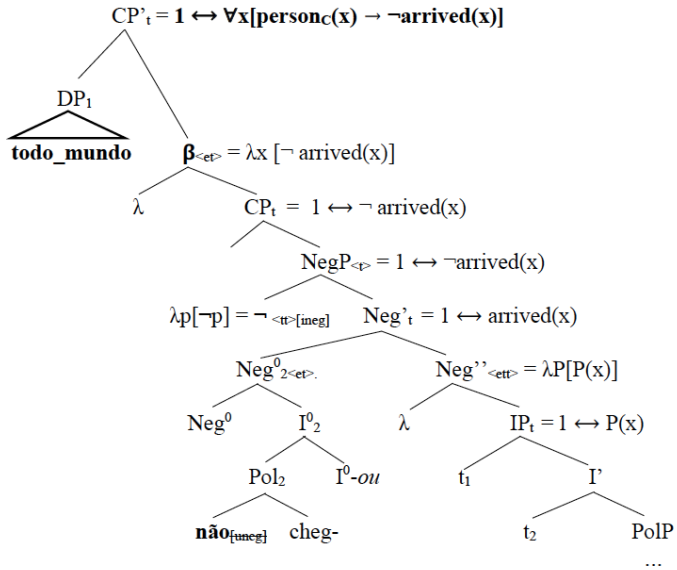
b.



Let us begin looking at the narrow scope reading of clauses (11a) and (12). Sentence (12) has the structure in (19):

- (19) a. $[_{CP} \text{ todo_mundo } [_{CP} [_{NegP} \neg_{[ineg]} [_{Neg'} [_{Neg^0} \text{ não}_{[ineg]} \text{ chegou}_2 [_{IP} t_1 t_2 [_{PolP} t_1 t_2]]]]]]]]]]$

b.

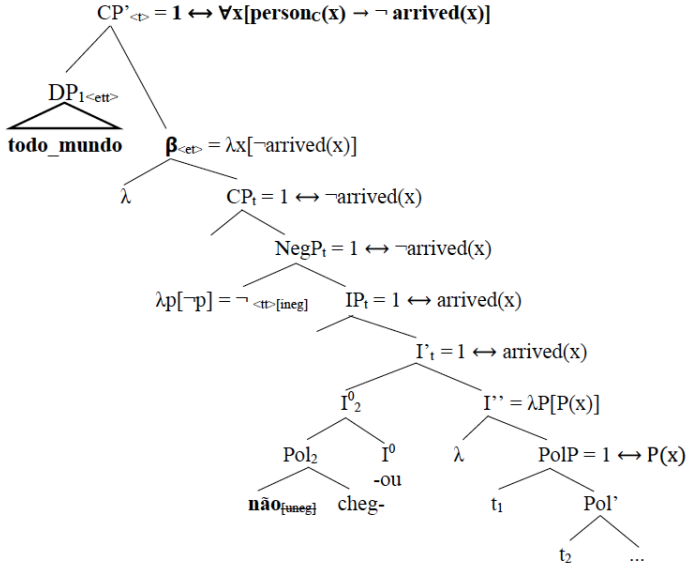


In the structure above, the universally quantified DP *todo mundo*, by QR, adjoins to CP, a position above the NEG operator. From this position, it c-commands NEG, and in virtue of this has the latter within its scope domain. As for *não*, it forms with V^0 and I^0 a complex head whose landing site is Neg^0 . Moreover, NEG c-commands the complex head $Neg^0 + I^0 + não + V^0$, and by doing so guarantees that *não* is in an agree relation with it, i.e., in a configuration where its uninterpretable feature can be checked by NEG. The outcome is the reading in which the abstract negation is under the scope of the universal quantifier.

As for the narrow scope reading of sentence (11a), it has the following structure:

(20) a. $[_{CP'} \text{todo_mundo}_1 [_{CP} [_{NegP} \neg_{[in\acute{e}g]} [_{Neg'} [_{IP} [_{I'} [_{I^0} \text{n\~{a}o}_{[u\acute{n}e\acute{g}]} \text{chegou}_2 [_{PolIP} t_1 t_2 [_{VP} t_1 t_2]]]]]]]]]]]$

b.

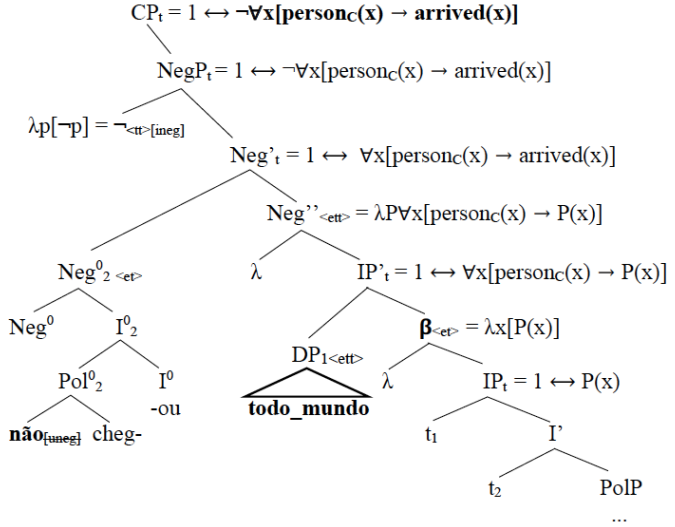


In (20), the GQ again adjoins to CP, i.e., above NegP, c-commanding the sentential negation. In turn, the semantically non-negative *nã*o, as a complex head with V^0 , raises to I^0 . Once again it is c-commanded by NEG. By satisfying Upward Agree, it gets its [uNEG] feature checked. As a result, the configuration in (20) gives rise to the truth conditions in which NEG is once more under the scope of the GQ.

Now, to generate the second available reading of (11a) and (12), i.e., the one in which negation scopes over the universal quantifier, it suffices to posit that the GQ *todo mundo* adjoins by QR to IP. When the complex head moves to Neg^0 , i.e., sentence (12), the result is the structure in (21). In this NEG c-commands both the complex head $Neg^0+I^0+não+V^0$ and the GQ. Thus, *nã*o is in an Upward Agree relation with the semantically negative operator, which has *todo mundo* in its scope.

- (21) a. $[CP [_{NegP} \neg_{[ineg]} [_{Neg'} [_{Neg''} [_{Neg^0} \text{n\~{a}o}_{[ineg]} \text{chegou}_2 [IP' \text{todo mundo}_1 [IP t_1 t_2 [_{PolIP} t_1 t_2 [_{VP} t_1 t_2]]]]]]]]]]]$

b.

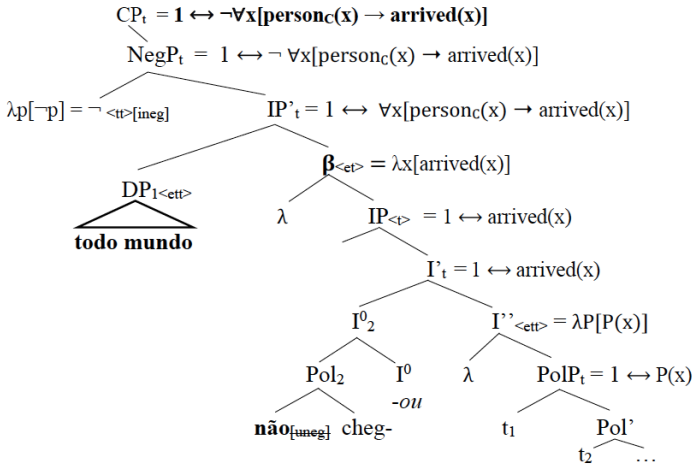


Therefore, the truth conditions assigned to (12), given (21), are the ones in which the abstract NEG operator scopes over the GQ.

Regarding (11a), it is easy to see that like in the narrow scope derivation the landing site of the complex head doesn't interfere at all in the scope interactions between NEG and the GQ. In (22), since *todo mundo* again adjoins by QR to IP, NEG c-commands it and $I^0 + n\tilde{a}o + V^0$. This structure again kills two birds with one stone: it allows the uninterpretable feature of the head to be checked against its interpretable counterpart hosted by NEG, and the sentential negation operator has again the GQ within its scope domain.

- (22) a. $[CP [_{NegP} \neg_{[ineg]} [_{Neg'} [IP' \text{todo_mundo}_1 [IP [I' [I^0 \text{n\~{a}o}_{[ineg]} \text{chegou}_2 [_{PolIP} t_1 t_2 [_{VP} t_1 t_2]]]]]]]]]]]$

b.



To sum up, the shared ambiguity of (11a) and (12) is due to the scope interactions between the phonologically null sentential NEG operator (type <tt>) in the Spec of NegP above IP and the GQ *todo mundo*. When *todo mundo*, by QR, adjoins to CP, i.e., to a position from where it c-commands the NEG operator, we get the reading 1 (i.e., $\forall > \text{NEG}$). On the other hand, by adjoining to IP, hence below NegP, *todo mundo* is c-commanded by the negation, deriving reading 2 (i.e., $\text{NEG} > \forall$). The semantically non-negative word *não*, whose presence is licensed by the abstract NEG operator, is just a manifestation of syntactic agreement. As the head of PolP above VP, it can, by integrating a complex head, occupy I^0 or raise to Neg^0 . In both cases it is in Upward Agree relation with the semantic negation and, therefore, gets its uninterpretable feature checked.

Thus, the analysis proposed in this section, by portraying negative sentences in BP that contain *não* as an instance of Negative Concord, provided a syntax and compositional semantics capable of assigning to (11a) and (12) the two readings they have. This is a clear advantage over a position that treats *não* as the semantic negation, since the latter falls short of accounting for the scope ambiguity exhibited by both constructions.

5 Conclusion

In this paper I argued that sentential negation in BP is an unpronounced NEG operator occupying the Spec of NegP above IP. I demonstrated that

the negative marker *não*, due to its syntactic status as a head that always forms a complex head with V^0 and I^0 via head movement, cannot have the semantics of a sentential negation. If it were the semantic negation, it would be a predicate negation. Consequently, the range of scope ambiguities in sentences containing *não* and universally quantified DPs is left unaddressed. In contrast, drawing on Zeijlstra's theory of Negative Concord (2004, 2008), I show that an analysis of *não* as a semantically non-negative element head of a PolP, which bears an uninterpretable [uNEG] in Upward Agree relation with an abstract sentential negation, accounts for the scope interaction between negation and universally quantified DPs in BP.

References

- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2001. Derivation by Phase. In Michael Kenstowicz (ed.), *Ken Hale: A Life in Language*. Cambridge, MA: MIT Press, 1–52.
- Heim, Irene & Angelika Kratzer. 1998. *Semantics in Generative Grammar*. Oxford, UK: Blackwell.
- von Fintel, Kai. 1994. *Restrictions on quantifier domains*. Doctoral dissertation, University of Massachusetts, Amherst.
- LaCara, Nicholas. 2016. Verb phrase movement as a window into head movement. *Proceedings of LSA* 1(17):1–14.
- Ladusaw, William. 1992. Expressing negation. In Chris Barker & David Dowty (eds.), *Proceedings of SALT* 2:237–259.
- Laka, Itziar. 1990. *Negation in syntax: On the nature of functional categories and projections*. Doctoral dissertation, MIT, Cambridge, MA.
- Lopes, Ruth & Maura Rocha. 2017. Adjunção. In Mary A. Kato & Milton do Nascimento (eds.), *Gramática do Português Culto falado no Brasil: volume 2: a construção da sentença*. São Paulo, Brazil: Contexto.
- May, Robert. 1977. *The grammar of quantification*. Doctoral dissertation, MIT, Cambridge, MA.

- May, Robert. 1985. *Logical Form: Its Structure and Derivation*. Cambridge, MA: MIT Press.
- Pollock, Jean-Yves. 1989. Verb Movement, Universal Grammar, and the Structure of IP. *Linguistic Inquiry* 20(3):365–424.
- Rizzi, Luigi. 1996. Residual verb second and the Wh-criterion. In Adriana Belletti & Luigi Rizzi (eds.), *Parameters and Functional Heads*. Oxford, UK: Oxford University Press, 63–90.
- Roberts, Ian. 2003. Head Movement. In Mark Baltin & Chris Collins (eds.), *The Handbook of Contemporary Syntactic Theory*. Malden, MA: Blackwell Publishing, 113–147.
- Schoorlemmer, Erik & Tanja Temmerman. 2012. Head Movement as a PF-Phenomenon: Evidence from Identity under Ellipsis. In Jaehoon Choi, E. Alan Hogue, Jeffrey Punske, Deniz Tat, Jessamyn Schertz, & Alex Trueman. (eds.), *Proceedings of the 29th West Coast Conference on Formal Linguistics*. Somerville, MA: Cascadilla Press, 232-240.
- Schwenter, Scott. 2005. The pragmatics of negation in Brazilian Portuguese. *Lingua* 115:1427–1456.
- Sousa, Lilian T. 2007. *Formas reduzidas de itens negativos no Português Brasileiro*. Master's Thesis. Universidade Federal de Minas Gerais, Brazil.
- Sousa, Lilian T. 2011. Sentential negation in Brazilian Portuguese: Pragmatics and syntax. *Journal Lipp* 1:89–103.
- Zeijlstra, Hedde. 2004. *Sentential negation and negative concord*. Doctoral Dissertation, University of Amsterdam. Utrecht, Netherlands: LOT Publications.
- Zeijlstra, Hedde. 2008. *Negative Concord is Syntactic Agreement*. URL: <https://ling.auf.net/lingbuzz/000645>
- Zeijlstra, Hedde. 2012. There is only one way to agree. *Linguistic Review* 29(3):491–539.

Ri-investigating inverse number in Dagaare*

DOUGLAS PULLEYBLANK, DARIUS ADJONG, ALEXANDER
ANGSONGNA, LISA MATTHEWSON, AND ARIWAN ADDY SUHAIRI
University of British Columbia

1 Introduction

Dagaare (Niger-Congo, Mabia/Gur; glottocode: sout2789) exhibits an interesting pattern of number marking whereby particular suffixal markers appear in the singular for one set of nouns, but in the plural for another set of nouns. This phenomenon is referred to as ‘inverse number marking’ (e.g., Corbett 2000:159), and is found in various languages around the world, including, for example, Kiowa (Kiowa-Tanoan; Watkins 1984). See Corbett (2000:159–166) for an overview.

Examples of inverse number marking from Dagaare are given in Table 1. We see that for the stems *bì* ‘child’ and *dò* ‘warthog/bush pig’, the suffix *-ri* appears on the plural form, but for the stems *kù* ‘tortoise’ and *kómbí* ‘tomato’, *-ri* appears on the singular form.¹

Table 1: Dagaare inverse number examples

Stem	Stem gloss	Singular	Plural
<i>bì</i>	‘child’	<i>bíé</i>	<i>bíi-ri</i>
<i>dò</i>	‘warthog’	<i>dùó</i>	<i>dò-ri</i>
<i>kù</i>	‘tortoise’	<i>kù-ri</i>	<i>kùé</i>
<i>kómbí</i>	‘tomato’	<i>kómbí-ri</i>	<i>kómbié</i>

* Hotze has always been a champion for students. His contribution to both the undergraduate and graduate programs at UBC has been consistent and massive. We therefore think this paper is a fitting tribute to his indefatigable work in this regard as it is a contribution of one former student, one visiting student, and one current student — oh, and it also drags one struggling phonologist into the scary domain of semantics... Thanks to Ryan Bochnak for very helpful comments on a draft of the paper. This paper was supported by a SSHRC Insight grant to Pulleyblank.

¹ The non-*ri* forms often have additional material, such as the final [é] in the singular of ‘child’ and the plural of ‘tortoise’ and ‘tomato’. Whether this additional material constitutes a morpheme or is a result of epenthesis is a matter we do not address here, since our focus is the *-ri* suffix. Grimm’s work assumes, following Anttila and Bodomo (2009), that these final vowels are epenthetic; Angsongna (2023) presents evidence against this assumption, suggesting that the vowel is a morpheme.

Grimm (2010, 2012a,b, 2018, 2021) proposes that the singular and plural uses of the same formatives can be explained by reference to ‘individuation’. The core of his proposal is that with nouns that are inherently singular and countable, *-ri* indicates multiple individuals, i.e., plurality. In nouns that are inherently plural and noncountable, *-ri* again indicates individuation, referring now to a (singular) component.

In this paper we test the individuation hypothesis against a database of Dagaare nouns. For each noun marked with *-ri* in either the singular or the plural, we independently evaluated whether the semantics suggests inherent individuation or not. We then investigated how the suffixal marking lines up with the semantics. At issue is whether the lexical semantics of nouns directly determines the choice of suffixes, or whether morphological encoding of suffix choice is necessary, with such encoding only correlating imperfectly with the lexical semantics. Our results support the latter conclusion.

2 A few notes on *-ri*

Before turning to our predictions and testing, it is appropriate to delineate the details of what we refer to as the suffix ‘*-ri*’. This suffix appears in a variety of surface forms. Consider the examples in (1) which illustrate, drawing on instances of plural *-ri*.

- (1) Variation in the surface forms of *-ri*
- | | | | |
|----|-----|---------|---------------------|
| a. | -rì | gbàg-rì | ‘agama lizard-PL’ |
| b. | -rí | láá-rí | ‘bowl-PL’ |
| c. | -rì | zú-rì | ‘head-PL’ |
| d. | -rí | kóg-rí | ‘chair-PL’ |
| e. | -nì | jén-nì | ‘sense-PL’ |
| f. | -ní | lón-ní | ‘frog-PL’ |
| g. | -nì | lón-nì | ‘hourglass.drum-PL’ |
| h. | -ní | gòn-ní | ‘silk.cotton-PL’ |

Three properties of the *-ri* suffix are predictable from the root. The tongue root value, [i] vs. [ɪ], is determined by root-controlled harmony. The tone, L vs. H, is determined by the root. The initial consonant is by default [r] but appears as [n] when the root contains a nasal consonant. See Anttila and Bodomo (2009) and Angsongna (2023) for details. We consider all eight surface realisations to be instances of the suffix *-ri*.

In addition, we might ask why singular and plural *-ri* are not simply two homophonous suffixes. In brief, there are two arguments against the

multi-suffix possibility. First, this hypothesis would lead one to expect the possibility of some nouns being marked for both singular and plural with *-ri*. We know of no instances where that happens. Second, there is an interesting idiosyncrasy illustrated by *-ri*, whether singular or plural. Nominal roots exhibit a complex pattern of length alternations. Some roots are consistently short, some are consistently long, and others alternate between short and long. For example, the root *gbó* ‘heart’ is consistently short: [gbô] ‘heart-SG’, [gbó-rì] ‘heart-PL’; the root *nú* ‘hand’ alternates between short and long: [nú] ‘hand-SG’, [núú-rì] ‘hand-PL’; the root *wóó* ‘elephant’ is consistently long: [wóó] ‘elephant-SG’, [wóó-rì] ‘elephant-PL’. Whether marking singular or plural, *-ri* consistently selects the long form of a root if there is one.

Overall, our assumptions about *-ri* being a single morpheme are consistent with those made by Grimm (2012a, 2021).

3 Predictions

Previous literature on inverse number marking often suggests that there is a semantic basis for the division into two sets of nouns (those for which the inverse marker indicates plural, and those for which it indicates singular). Corbett (2000:162) notes that there is a “notion of an inverse marker which indicates the less expected number.” Corbett also observes (2000:161) that “the two main classes of noun in Kiowa, one with the inverse marker for plural and the other with inverse marking for singular, conform broadly with the Animacy Hierarchy, since the first contains all the animates.”

Grimm (2021:454) argues that “the inverse number system in Dagaare reflects principled lexical semantic categorization”, although he considers that the system no longer applies to newly created or imported nouns, and he also notes that that “there is a certain amount of conventionalization, historical residue and fuzzy boundaries in the Dagaare system” (2021:455).

Grimm offers two generalizations about the semantic underpinning of the Dagaare inverse number system: frequency and individuation. Frequency refers to the claim that “when a noun designates an entity which is likely to appear singly, *-ri* encodes the plural, while when a noun designates an entity which is likely to appear in multiples, *-ri* encodes the singular” (2021:453). This proposal is in line with Corbett’s idea about the “less expected” form being the overtly marked one. Individuation refers to the distinction between referents which are conceived of as

individuals and those that are conceived of as “a collection of entities or an amorphous mass” (Grimm 2021:457).

Grimm (2021:451;457; see also Grimm 2018) argues that Dagaare divides its nouns into four individuation categories. In increasing degrees of individuation, these denote liquids and substances, granular aggregates, collective aggregates, and individuals. The first two of these categories are claimed not to appear with *-ri*. The collective aggregate nouns are termed “basic plural” and take *-ri* in the singular, and the individual nouns are termed “basic singular” and take *-ri* in the plural. Grimm (2012a) also outlines a set of more specific predictions, quoted in (2).

- (2) Grimm’s (2012a:83) predictions for *-ri*-marking
- i. Nouns for higher-level (more salient) animals are more likely to be unmarked in the singular than nouns for insects (animacy)
 - ii. Nouns for trees should be unmarked in the singular in comparison to nouns for vegetation (distinguishability)
 - iii. Nouns for tools should be more likely to be unmarked in the singular than the converse (one canonically interacts with them individually)
 - iv. Nouns for body parts which inherently come in pairs or groups should be more likely to be unmarked in the plural than not, while nouns for body parts which inherently come in single units should be more likely to be unmarked in the singular than not

In our study we set out to test these predictions, using a database of forms described in the next section.

4 Methodology

The data used in this paper are from the central variety of Dagaare spoken in Sombo in the Nadowli-Kaleo district, Ghana. The data were collected in Ghana from twenty-three (23) native speakers in the months of March and April 2018. It involved the elicitation of wordlists, phrases, and sentences and was based on the SIL Comparative African Wordlist (Snider & Roberts 2004). Short stories, songs, and descriptions of local events/culture also formed part of the database. This was supplemented by data from prior literature and data from one of the authors, Alexander Angsongna, who is a native speaker of the above variety of central

Dagaare; the supplemental data served to fill in gaps where a singular or a plural appeared in our data collected in Ghana, but not both. The elicitation was done with a Shure WH30XLR cardioid condenser (a headset microphone) and Rode NGT2 supercardioid condenser (a shotgun microphone) at the sampling rate of 48 kHz and bit depth of 16 bits. The microphones were attached to a Zoom Q8 camera.

The steps towards arriving at the results in this paper involved a number of stages. We started with a total database of seventeen thousand three hundred and fifty-nine (17,359) entries. These entries included duplications, verbs, nouns, adjectives, particles such as tense, negation, and focus particles. Since our focus is on nouns, the next step was to separate the nouns from the rest of the database. We did this using Microsoft Excel.

This stage resulted in a total of four thousand one hundred and seventy-two (4,172) nouns. With this number, we took some further steps. We removed all compound nouns (except for cases where the second member of the compound did not appear independently in the database). We eliminated derived nouns which resulted from nominalization and reduplication. Incorrect entries were also removed. Where stems, including loan words, did not have a clear marker of number or where the alternative was not in the database, we were able in certain instances to edit the entry to include the missing singular or plural form; in other cases, the entry was removed. In terms of number morphology, zero suffixes had not been systematically glossed; so, we added glosses where relevant. Moreover, if a tonal or segmental error was noticed while checking an entry, it was corrected, though we did not systematically try finding such errors for all entries. We also edited cases where glossing was unusual.

After completion of the above steps, we arrived at a total of four hundred and thirty-one (431) simple nouns.² The nouns were grouped based on shared roots and we ensured that identical words were adjacent to each other. Out of the 431 nouns, some had both singular and plural forms; some had only singular forms and some had only plural forms. Data gaps were flagged and filled in by Alexander Angsongna.

Our main research objective involves determining whether a noun root is intrinsically individuated or not (or identifiable by some semantic criterion — see Grimm 2010, 2012a, 2018) and whether that determines

² By ‘simple noun’, we refer to a noun root + number suffix combination. That is, all simple nouns are two-morpheme sequences (root+suffix, though the suffix can in some cases be \emptyset).

the choice of number affixes. So, with regards to semantic assessment, two spreadsheet files were created with the 431 nouns: one with information about the affix choice and the other with information about roots. Each of the authors semantically coded the roots independently, and without consulting the file that contained the affix choice information. We then met to compile the individual assessments into a group consensus. Assessments included mass vs. count, individuated vs. grouped, perceivable components vs. no perceivable components; we also did an assessment into semantic categories — human, body parts, animal, food, insect, event, etc. Based on our combined assessments, we were able to decide on semantic descriptions and categories for each noun. See Section 5 for detailed semantic descriptions and classification of nouns. When semantic coding was complete, we combined the data with our classification of roots to test the correlations of semantics with morphological singular-plural choice. We report here on affix choice for the 222 nouns in our database that used *-ri* in either the singular or the plural.

5 Results

We present our results here by first expressing the precise prediction being tested and then giving the counts from our database. We classified nouns according to three properties: (i) count vs. mass; (ii) if count, then individual vs. group; (iii) if mass, then perceivable components vs. no perceivable components. This corresponds to Grimm’s four individuation categories as laid out in Table 2.

Table 2: Individuation categories

Grimm’s categories	Count vs. mass	Individual vs. group	Perceiv. vs. no perceiv.
Liquids; substances	mass		no perceivable components
Granular aggregates	mass		perceivable components
Collective aggregates	count	group	
Individuals	count	individual	

Our predictions based on these individuation categories are laid out below.

First, regarding (i) liquids and substances and (ii) granular aggregates, we may distinguish strong and weak predictions. The strong prediction is that there should be no use of *-ri* as a suffix for these two classes of nouns. In Table 3, we present our results. The number of cases involving *-ri* for each of the four individuation classes is compared with the total number of examples in that class.

Table 3: Occurrence of *-ri* in the four individuation classes, based on 449 nouns

	Number of <i>-ri</i>	Total number in class
Liquids; substances	21	49
Granular aggregates	3	15
Collective aggregates	48	101
Individuals	150	284

As shown in Table 3, there are examples of *-ri* in both the liquids and substances class and the granular aggregates class, inconsistent with the strong prediction. While there are only a few examples of *-ri* with granular aggregates, over 40% of the liquid and substances class is marked by *-ri*.

The weak prediction for these two classes is that if *-ri* is found for nouns of these types, then it should mark the singular, not the plural. Our results are given in Table 4.

Table 4: Occurrence of *ri*-singular and *ri*-plural in mass nouns, based on 24 nouns with *-ri*

	Prediction	<i>ri</i> -singular	<i>ri</i> -plural
Liquids; substances	ri-SG > ri-PL	11	10
Granular aggregates	ri-SG > ri-PL	2	1

As seen, even the weak prediction is not met. The number of nouns using *-ri* in the singular in the two mass noun classes is comparable to the number of nouns using *-ri* in the plural.

Regarding count nouns, *-ri* is expected to occur with both collective aggregates and individuals. As seen in Table 3, this is indeed the case: approximately half of both count classes have number marking with *-ri*. The more important prediction for these classes — the core of the inverse

numbering pattern — is that *-ri* should specifically occur with singulars in the collective aggregate class and with plurals in the individuals class. Our results are given in Table 5.

Table 5: Occurrence of *ri*-singular and *ri*-plural in count nouns, based on 198 nouns with *-ri*

	Prediction	<i>ri</i> -singular	<i>ri</i> -plural
Collective aggregates	ri-SG > ri-PL	23	25
Individuals	ri-SG < ri-PL	45	105

Our results are not consistent with the individuation prediction. The number of instances of *-ri* in singular collective aggregates is comparable to the number in the plural. For the individuals class, the predicted asymmetry holds as a weak tendency, but roughly a third of this class occurs with *-ri* in the singular, against expectation.

Consider next the predictions for nouns of particular semantic types. In terms of animals and plants, it is predicted that for higher-level animals, the default interpretation would be singular, so *-ri* is more likely to occur in the plural, while for lower-level animals such as insects, the default would be plural, so *-ri* is more likely to occur in the singular. In the class of ‘animals’, we included mammals and reptiles; fish and birds were not included in our counts as we were unsure how to control for interpretations involving schooling or flocking. In a similar vein, trees are more likely to have a default singular interpretation and therefore take *-ri* in the plural while less distinguished vegetation is more likely to have a default plural interpretation and therefore take *-ri* in the singular (Grimm 2021:452).

Table 6: Occurrence of *ri*-singular and *ri*-plural in animals and plants

	Prediction	<i>ri</i> -singular	<i>ri</i> -plural
Animals	ri-SG < ri-PL	5	23
Insects	ri-SG > ri-PL	3	8
Trees	ri-SG < ri-PL	0	6
Vegetation	ri-SG > ri-PL	3	5

Our results are consistent with the predictions when *ri*-singulars are predicted to be fewer in number than *ri*-plurals, and inconsistent with the predictions when a greater number of singular forms are expected. The

overall tendency is simply for *ri*-plurals to outnumber *ri*-singulars in these semantic classes.

For tools, the default is expected to be singular; hence the prediction for *-ri* is that it should occur in the plural.

Table 7: Occurrence of *ri*-singular and *ri*-plural in tools

	prediction	<i>ri</i> -singular	<i>ri</i> -plural
Tools	ri-SG < ri-PL	10	17

The prediction is confirmed as a tendency only, with over a third of tools that are marked for *-ri* taking *-ri* in the singular.

Finally, we considered the use of *-ri* in nouns denoting body parts. As sketched in (2), it is predicted that body parts that are grouped would have a default plural interpretation (hence *-ri* in the singular) while body parts that are not grouped would have a default singular interpretation (hence *-ri* in the plural) (Grimm 2021:453). The results are given in Table 8.

Table 8: Occurrence of *ri*-singular and *ri*-plural in body parts

	Prediction	<i>ri</i> -singular	<i>ri</i> -plural
Paired/grouped body parts	ri-SG > ri-PL	14	6
Unpaired/ungrouped body parts	ri-SG < ri-PL	14	15

As seen, individual body parts are quite evenly distributed between *ri*-singular and *ri*-plural — inconsistent with the prediction. Paired/grouped body parts are consistent with the prediction as a tendency, though we find six examples of ‘default’ singular.

6 Discussion

6.1 Semantics of *-ri*

We have shown so far that *-ri* is a single morpheme, which sometimes conveys singularity and sometimes plurality, and as seen in our results section, which of these meanings *-ri* conveys is not predictable from the semantics of the noun to which it attaches. It is a non-trivial challenge to find a unified semantic denotation for *-ri* that achieves the apparently

opposite results of sometimes conveying singularity and sometimes plurality.

Grimm (2012a) considers two proposals for the semantic analysis of *-ri*. The first analysis, which he adopts (see also Grimm 2021), assumes an *exclusive* interpretation of the plural, according to which plural denotations exclude singular atoms (e.g., Link 1983). Thus, an exclusive plural noun ‘children’ would refer only to pluralities of children, and give rise to falsity when applied to a single child. The function of *-ri* is then to select the complement set of the denotation of the noun: *-ri* added to a basic singular produces a plural noun that denotes only sums, while *-ri* added to a basic plural produces a singular noun that denotes only atoms.

The second analysis, which Grimm considers but then rejects, assumes an *inclusive* interpretation of *-ri* plurals, according to which these plural denotations include both sums and atoms (e.g., Krifka 1989). The contribution of *-ri* under this analysis is to produce closure under join, which means that a *-ri*-noun, regardless of whether *-ri* combines with a basic singular or a basic plural, will always denote the entire semi-lattice (covering both sums and atoms). The fact that basic singulars + *-ri* denote pluralities is achieved by pragmatic blocking. The same explanation can also account for why basic plurals + *-ri* denote singularities, under the assumption that the basic plural nouns have exclusive plural denotations, which are then made into inclusive plurals (containing atoms) by *-ri* (Grimm 2012a:96).

It seems to us that neither of these two analyses quite works. The exclusive plural analysis fails because the facts do not support an exclusive interpretation for *-ri*-plurals. This is illustrated in (3) to (4). The exclusive plural analysis predicts that answer B’ in (3) will be felicitous, since the plural form *bíírí* denotes only non-atomic sums and, therefore, if B has one child, it will be appropriate to deny that they have *bíírí*. This prediction does not fit the judgments of Alexander Angsongna.³

- (3) A: fò táá ná⁴ bìi-ri
 2SG have FOC child-RI
 ‘Do you have children?’

³ Grimm twice alludes to the fact that inclusive plural tests yield parallel results in Dagaare to in English (which has inclusive plurals) (2012a:96–97); he nevertheless opts for the opposite analysis. The only data he provides to test the inclusivity of plurals do not include *-ri*, hence is not a relevant example (2012a:97).

⁴ Note that the focus particle as indicated here has another variant referred to as *lá* in other varieties of Central Dagaare especially the Jirapa dialect. It also has clitic forms as *-ŋ*, *-e/-ε*.

B: Mím, Ñ táá ná bì-jèni
 yes 1SG have FOC child-one
 ‘Yes, I have one child.’

B’:# Ààjí, Ñ táá ná bì-jèni
 no 1SG have FOC child-one
 ‘No, I have one child.’

Example (4) makes a similar point. The negation of a *-ri*-plural negates both sums and atoms, not merely sums, as shown by the fact that C’s utterance conveys that C has not even one child.

(4) *Context: C has one child. C tells D:*

Ñ bá tàà bíí-rí
 1SG NEG have child-RI
 ‘I don’t have children.’

On the other hand, the inclusive plural analysis relies on the assumption that the simple basic-plural nouns have exclusive plural denotations, as noted above. These are then converted to inclusive plural denotations by *-ri*, and pragmatic blocking by the exclusive-plural bare noun results in a singular denotation for the *-ri* form. As noted by Grimm, however, this also does not fit the facts for simple plural nouns. This is illustrated in (5).

(5) Q: fù táá ná kòm̀bi-è
 2SG have FOC tomato-PL⁵
 ‘Do you have tomatoes?’

A: mím, Ñ táá ná kóm̀bi-yèni
 yes 1SG have FOC tomato-one⁶
 ‘Yes, I have one tomato.’

⁵ Concerning glossing, we have glossed [kòm̀bi-è] as ‘tomato-PL’ here, in line with an analysis of -è as a morpheme; if -è is an epenthetic vowel then the glossing would be more appropriately ‘tomato.PL’. See discussion of what is at stake in Section 6.2.

⁶ As pointed out by Ryan Bochnak, the semantic function of *yèni* in this example raises interesting questions. If ‘tomato’ is plural-denoting by default, then what exactly is the effect of adding *yèni* ‘one’? This and other questions concerning number in Dagaare go beyond our examination of *-ri* in this paper and require future research.

While we do not have a worked-out formal solution to this problem at the current time, the desired effect of *-ri*, based on the data in (3) to (5), is clear: when *-ri* applies to a simple singular denotation that includes only atoms, it *adds sums* to result in an inclusive plural denotation. When *-ri* applies to a simple plural denotation that is inclusive (containing both atoms and sums), it *removes the sums* to result in a singular denotation. Crucially, whether a simple noun will denote singularities or pluralities is *not* predictable from whether the noun refers to items that are typically found in groups or singularities in the real world. That is, we assume that it must be lexically specified for each simple noun whether it is inherently singular or inherently plural.

6.2 Lexical encoding of number

As discussed above (in Section 6.1), under Grimm's analysis the suffix *-ri* does not itself denote the singular or plural, but instead denotes negation of the lexical denotation of the base. Nouns may be lexically singular (e.g., *bi* 'child') or plural (e.g., *kù* 'tortoise'). For Grimm, lexical number-marking is systematically determined by the degree of individuation: more individuated nouns are predicted to be lexically singular.

Rejecting the inverse-marker analysis: If we reject entirely the analysis of *-ri* as an inverse marker, then there must be two distinct but homophonous morphemes *-ri*_[PL] and *-ri*_[SG], explaining how the "same" morpheme can mark either the singular or plural depending on the noun base that it attaches to. The choice of *-ri*_[PL] or *-ri*_[SG] is an idiosyncratic property of a given noun (within one of the noun classes that surfaces with a *-ri* suffix), and is presumably encoded in its lexical entry. However, the uniform morphophonological behaviour of the *-ri* suffix(es), as discussed in Section 2, strongly suggests that *-ri* is indeed a single morpheme, and this generalisation is lost if we postulate multiple homophonous *-ri* suffixes.

Rejecting the individuation analysis: However, it may be possible to reject Grimm's analysis of noun individuation in Dagaare while retaining the insight that *-ri* is an inverse marker. Suppose that nouns are arbitrarily specified in the lexicon as either [singular] (denoting atoms) or [plural] (denoting atoms + pluralities), rather than basic number being determined by a putative scale of individuation. Since *-ri* is an inverse marker and denotes the negation of the lexical base, we derive the observed pattern of number marking.

In general, however, analysing *-ri* as an inverse marker leaves behind the uneasy residue of the additional *-V* that appears in the “unmarked” forms of many nouns. As mentioned in footnote 2, whether or not *-V* is an epenthetic segment or a morpheme is unclear. While Grimm follows Anttila and Bodomo (2009) in assuming that its presence is phonologically conditioned and thus predictable, Angsongna (2023) raises problems for the epenthesis account. Future research is needed to determine the optimal analysis of this marker. If *-V* can be shown definitively to *not* just be an epenthetic segment, then the inverse-marker analysis must also provide an account for *-V*.

6.3 Borrowings

Grimm (2021:454) states that inverse number is not observed in loan vocabulary. Particularly since we have argued above that inverse number is not (fully) predictable even in native vocabulary, this would certainly not be surprising. Nevertheless, consider borrowed items such as those in (6) and (7). The forms in (6) appear to be phonologically special, as noted in Grimm (2021), since the *-ri* observed in the singular could be the Dagaare interpretation of the phonological form of the English.

(6) Singular marked by *-ri*

Singular	Plural	Source	Gloss
lóó-ri	lóè	English	‘lorry’
sákì-ri	sákìè	English	‘bicycle’
hántǽ-ri	hántǽè	English	‘handkerchief’

The forms in (7), which do not show such a phonologically motivated effect, might be interpreted as showing a general tendency to use *-ri* to mark plural in loan words. It is noteworthy, however, that all the nouns here can be seen as designating entities that are likely to occur singly, which would lead us to expect by the inverse number hypothesis that *-ri* should encode plural.

(7) Plural marked by *-ri*

Singular	Plural	Source	Gloss
bìríǽ	bìríǽ-rí	English	‘brick’
wáǽ	wáǽ-rí	English	‘watch’
dókítà	dókítà-rí	English	‘doctor’
kópò	kópò-rí	English	‘cup’

mónḡò	mónḡò-rí	English	‘mango’
bilédi	bilédi-rí	English	‘blade’
ásíbítì	ásíbítì-rí	English	‘hospital’
néésì	néésì-rí	English	‘nurse’
pèrikó/pòrikó	pèrikó-rí	Akan/Portuguese	‘pig’
kòdú	kòdú-rí	Akan	‘banana’
pòlìsì	pòlìsì-rí	English	‘police’
sùkúù/sàkúù	sùkúù-rì/sàkúù-rì	English	‘school’
kàníè	kàní-rì	Akan	‘lantern’
góótà	góótà-rí	English	‘gutter’
táájà	táájà-rí	English	‘tyre’
bókítì	bókítì-rí	English	‘bucket’
sódzà	sódzà-rí	English	‘soldier’
ǰěnsì	ǰěnsì-rí	English	‘sheet’

We leave an investigation of the productivity of *ri*-singular vs. *ri*-plural for future investigation.

6.4 Comparative Mabia

Aside from central Dagaare, the morpheme *-ri* has a number-marking function in three other dialects of Dagaare/Dagara, namely Lobr, Wiile, and Birifor (Mwinlaaru 2023). A similar singular–plural alternation or inverse marking strategy involving *-ri* is found in these three dialects. Also, as in Dagaare/Dagara, a *-ri* morpheme is employed as a number marker in other Mabia/Gur languages. Some of these languages, e.g. Buli (Schwarz 2005, 2012; Akanlig-Pare 2005), Konni (Cahill 1999), and Moore (Delplanque 1995), employ *-ri* primarily as a singular, but Dagbani (Wilson 1972; Olawsky 1999), another Mabia/Gur language, employs *-ri* as a plural marker. A few other Mabia/Gur languages (e.g., Gurene – Dakubu 1996; Nsoh 2002) have no *-ri* for number marking.

On phonological grounds, it appears that *-ri* is not the original number marker in Mabia/Gur languages. One piece of evidence can be deduced from the distributional features of /r/. The approximant /r/ rarely occurs as a word-initial element in Mabia/Gur languages. In word-medial position, where /r/ is found frequently, /r/ occurs as an allophonic variant of a different sound. Dagaare has several allomorphs of the *-ri* suffix (Angsongna 2023). It thus appears that there was an original morpheme (not *-ri*) in Dagaare and other dialects of Dagara that has etymologically been replaced by [-ri]. In Buli, [ri] and [di] are singular-marking variants in nouns like [bìisírì]~[bìisídí] ‘breast’, [nísírì]~[nísídí] ‘hand’, and

[nùènsírí]~ [nùènsídí] ‘footwear’. The Buli pronominal system provides good evidence for [di] being the original form of the morpheme. For example, in accordance with the agreement pattern of Buli, morphemes that mark number also function as independent pronouns. All nouns that have [-ri] or [-di] as a singular suffix in Buli select [di] as their independent pronoun; [ri] never occurs as a pronoun. There are also Mabilia languages that have [-di] but no [-ri] as a number marker, so we would assume that not every Mabilia language has developed a [-ri] variant.

As seen above, most of the Mabilia languages about which we have discussion take *-ri* as a marker of singular. This suggests that the proto-language had *-ri* as a singular morpheme.⁷ If this is correct, then we would expect the innovation in Dagaare to be the use of *-ri* as a plural marker. Taken together with the individuation hypothesis concerning the semantics of *-ri*, we would expect that plural cases involving *-ri* would be more semantically coherent than singular cases involving *-ri*. This follows since the cases in the proto-language with *-ri* as a singular would not be expected to show individuation distinctions: *-ri* simply marks singular. As *-ri* shifted to encoding plurality on certain nouns, if the individuation hypothesis is correct, then we would expect a change only in count nouns where the default meaning is individuals. That is, liquids/substances, granular aggregates, and collective aggregates would be expected to continue using *-ri* in the singular since there would be no pressure for change.

This is easy to test. There are 81 nouns that show *-ri* in the singular and 141 nouns that show *-ri* in the plural. The breakdown in terms of the four individuation categories we have been considering is shown in Table 9.

While the effect is not absolute, we see that count/individual cases where *-ri* appears in the plural constitute 74% of all *ri*-plural forms while only 56% of all *ri*-singular forms. Overall, the *ri*-singular nouns are indeed more semantically diverse than are the *ri*-plural forms.

⁷ This hypothesis leaves unexplained the Dagbani pattern where *-ri* is a marker of plural only.

Table 9: Dagaare individuation in *ri*-singular and *ri*-plural

	Nouns taking ri-SG		Nouns taking ri-PL	
		Percentage		Percentage
Liquids; substances	11	14%	10	7%
Granular aggregates	2	2%	1	1%
Collective aggregates	23	28%	25	18%
Individuals	45	56%	105	74%

7 Conclusion

In this study, we tested semantic individuation as a means of determining the use of *-ri* in Dagaare as a singular or plural marker. Using a database of forms collected from multiple speakers, we coded nouns for individuation categories and assessed these categories for observed use of singular and plural *-ri*. While we did not find consistent enough use of semantic individuation to directly predict observed suffix choice, we did find certain indications that Dagaare has been innovating in the direction of including *-ri* as a plural marker, and doing so preferentially for nouns denoting countable individuals.

We leave numerous questions for future investigation. Notably, we have not developed a formal treatment of the semantics of *-ri* suffixation. In addition, we have not considered how to integrate the properties of *-ri* into a larger treatment of number in Dagaare generally. In particular, we have not addressed the morphology and semantics of the epenthetic/lexically specified vowel ‘suffix’ that is paired with *-ri*, and we have not considered the suffixes that account for the roughly 50% of the lexicon that marks number in other ways.

References

- Akanlig-Pare, George. 2005. *Buli Tonology: a non-linear approach*. Ph.D. dissertation, University of Ghana, Legon.
- Angsongna, Alexander. 2023. *Aspects of the morphophonology of Dagaare*. Ph.D. dissertation, University of British Columbia, Vancouver, BC.
- Anttila, Arto & Adams Bodomo. 2009. Prosodic morphology in Dagaare. In Masangu Matondo, Fiona Mc Laughlin, & Eric Potsdam (eds.), *Selected Proceedings of the 38th Annual Conference on African Linguistics (ACAL 38)*. Somerville, MA: Cascadilla Press, 56-68.

- Cahill, Michael C. 1999. *Aspects of the Morphology and Phonology of Kɔnni*. Ph.D. dissertation, Ohio State University, Columbus, OH.
- Corbett, Greville G. 2000. *Number*. Cambridge, UK: Cambridge University Press.
- Dakubu, M. E. 1996. *Grammar of Gurune*. Trial Edition. Language Centre, Legon.
- Delplanque, Alain. 1995. Que signifient les classes nominales? L'exemple du mooré, langue gur. *Linguistique africaine* 15:5–56.
- Grimm, Scott. 2010. Number and markedness: A view from Dagaare. *Proceedings of Sinn und Bedeutung* 14:168–184.
- Grimm, Scott. 2012a. Individuation and inverse number marking in Dagaare. In Diane Massam (ed.), *Count and mass across languages*. Oxford, UK: Oxford University Press, 75–98.
- Grimm, Scott. 2012b. *Number and Individuation*. Ph.D. dissertation, Stanford University, Stanford, CA.
- Grimm, Scott. 2018. Grammatical number and the scale of individuation. *Language* 94(3):527–574.
- Grimm, Scott. 2021. Inverse number in Dagaare. In Patricia C. Hofherr and Jenny Doetjes (eds.), *The Oxford handbook of grammatical number*. Oxford, UK: Oxford University Press, 445–462.
- Krifka, Manfred. 1989. Nominal reference, temporal constitution and quantification in event semantics. In Johan van Benthem, Renate Bartsch, & P. von Emde Boas (eds.), *Semantics and contextual expression*. Dordrecht, Netherlands: Foris Publications, 75–115.
- Link, Godehard. 1983. The logical analysis of plurals and mass terms: A lattice-theoretical approach. In Rainer Bäuerle, Christoph Schwarze, & Arnim von Stechow (eds.), *Meaning, use, and interpretation of language*. Berlin, Germany: De Gruyter.
- Nsoh, Avea. 2002. Classifying the nominal in the Gurene dialect of Farefare of Northern Ghana. *Journal of Dagaare Studies* 2:83–95.
- Mwinlaaru, Isaac N. 2023. *What really counts in nominal classification in Dagaare: A Mabilia language*. Available at SSRN 4347091.
- Olawsky, K. J. 1999. *Aspects of Dagbani grammar: with special emphasis on phonology and morphology*. München, Germany: Lincom Europa.

- Schwarz, Anne. 2005. *Aspekte der Morphosyntax und Tonologie im Buli (mit Schwerpunkt auf dem Buli von Wiaga)*. Doctoral dissertation, Humboldt Universität, Berlin.
- Schwarz, Anne. 2012. *A1. Buli*. In Gudrun Mieke, Brigitte Reineke, & Kerstin Winkelmann (eds.), *Noun class systems in Gur languages: North Central Gur languages, vol. 2*. Köln, Germany: Rüdiger Köppe Verlag, 39–61.
- Snider, Keith & James Roberts. 2004. SIL comparative African wordlist (SILCAWL). *Journal of West African Languages* 31(2):73–122.
- Watkins, Laurel J. 1984. *A grammar of Kiowa*. Lincoln, NE: University of Nebraska Press.
- Wilson, William. A. A. 1972. *Dagbani: An introductory course*. Tamale: GILLBT.

Tell you what. English has quexistentials.*

DANIEL K. E. REISINGER
University of British Columbia

1 Introduction

A common claim in the recent literature on quexistentials is that English lacks forms that can function *both* as question words *and* as existential indefinites (e.g., Roelofsen et al. 2019; Hengeveld et al. 2023). In this regard, English seems to differ from other languages which have such forms. Consider, for instance, German *was* and Dutch *wat* which can either mean ‘what’ or ‘something’, as shown in (1) and (2). In English, on the other hand, the Wh-word *what* is usually restricted to the interrogative reading (3a), while the indefinite reading must be realized with a different lexical item, namely *something* (3b).¹

(1) *German:*

- a. **Was** hat Saoirse gehört?
QUEx has Saoirse heard
‘**What** did Saoirse hear?’
- b. Saoirse hat **was** gehört.
Saoirse has QUEx heard
‘Saoirse heard **something**.’

* In September 2016, when I started my journey at UBC, Hotze was the first faculty member that approached me. Over a cup of coffee at the infamous *Bean around the World*, we immediately bonded over our shared history at the University of Alberta and a profound fascination with semantics. I felt privileged to have found such a knowledgeable, kind, and witty mentor. Many years have passed since then, filled with research projects, co-authored papers, and the occasional pandemic, but as for Hotze, not much has changed. He is still as knowledgeable, kind, and witty as he was back then, and I have nothing but affection for him and his guidance. *Mögest Du auf ewig jung bleiben!*

¹ Some English indefinites, particularly *somewhat*, *somehow*, and *somewhere*, are clearly derived from Wh-words. However, since these forms also involve additional overt morphology, namely the presence of *some-*, they tend not to be treated as “proper” quexistentials (e.g., Hengeveld et al. 2023).

(2) *Dutch:*

- a. **Wat** heeft Saoirse gehoord?
QUEx has Saoirse heard
'**What** did Saoirse hear?'
- b. Saoirse heeft **wat** gehoord.
Saoirse has QUEx heard
'Saoirse heard **something**.'

(3) *English:*

- a. **What** did Saoirse hear?
- b. Saoirse heard **something** / **#what**.

While this pattern holds for English by and large, I will argue that the absence of quexistentials in the language is not absolute. At least in a small set of idiomatic expressions, Wh-indefinites seem to have found a niche. Consider, for instance, the bolded constructions in (4) and (5).

- (4) Oh, now, listen. **I tell you what**. I have an idea. Let me finish this while you go home and have a long hot bath, and I'll call round, we'll have dinner later, okay?

(Bridget Jones's Diary [2001 film])

- (5) **You know what?** I just realized. That's my journal. I bought it at a bookstore down the street.

(Jody Elizabeth Gehrman: Notes from the Backseat)

In these constructions, the Wh-word *what* obviously does not serve as an interrogative but rather as an indefinite. This can be highlighted by substituting *what* with *something*, as in (6) and (7). The interpretation of the utterance remains unaffected by this substitution.²

² Some speakers might prefer the use of the *will* future in (6), i.e., *I'll tell you something...*, though a look at example sentences from the *Corpus of Contemporary American English (COCA)* suggests that the version without *will* is in use as well, as exemplified in (i):

- (i) Well, **I tell you something**. The crew were fabulous, fabulous. So don't say anything wrong with the crew. They really did a great, great job.

(CNN: "Sick at Sea" [2003])

- (6) Oh, now, listen. **I tell you something.** I have an idea. Let me finish this while you go home and have a long hot bath, and I'll call round, we'll have dinner later, okay?
- (7) **You know something?** I just realized. That's my journal. I bought it at a bookstore down the street.

Occasionally, both types of indefinites may even occur together, as in (8) and (9), highlighting that *what* and *something* serve a similar function in these constructions.

- (8) **I'll tell you what. I'll tell you something.** My friends, if I had to lose Jonny to anyone, I can't imagine a more perfect woman than Halley. *(Serendipity* [2001 film])
- (9) **You know what? You know something? You know something?** If you had told us one year ago that we were going to come in third in Iowa, we would have given anything for that.
(The New York Times: "Howard Dean's Remarks to His Supporters" [2004-01-19])

The remainder of this squib is dedicated to these two intriguing quexistential constructions. First, in Section 2, I will show how we can distinguish the indefinite *tell you what* and *you know what* constructions from other English utterances which look similar on the surface but pattern quite differently in certain crucial ways. Section 3 will examine the form and function of the *tell you what* construction, while Section 4 will do the same for the *you know what* construction. Once this has been done, Section 5 will present a short diachronic corpus survey, highlighting that both constructions do not represent a recent innovation but have — in some form or other — been in use for several hundred years. A short summary in Section 6 concludes this investigation.

2 Similar constructions

Before we examine the indefinite *tell you what* and *you know what* constructions in detail, it is necessary to distinguish them from some other English constructions which look similar on the surface but behave quite differently in certain respects: ellipsis constructions and echo question constructions. Consider, for instance, the utterances in (10) to (12):

(10) *The indefinite construction:*

- a. Oh, now, listen. **I tell you ↘what**. I have an idea. Let me finish this [...] (*Bridget Jones's Diary* [2001 film])
- b. **You know ↗what?** I just realized. That's my journal. [...] (Jody Elizabeth Gehrman: *Notes from the Backseat*)

(11) *The ellipsis construction:*

- a. What do you think President Trump had to do with it? **I'll TELL you ↘what**. Like, how about everything?
(NBC News: "Today" [2018-04-29])
- b. NUDIE: Shit. What is this? Half rot-gut?
HANK: What are you talking about?
NUDIE: **You KNOW ↘what**. This here's more booze than soda.
(*Eight Scenes from the Life of Hank Williams* [1990 film])

(12) *The echo question construction:*

- a. A: ... and then you will tell me that you love me.
B: **I will tell you ↗WHAT?!**
- b. A: I know that the priest is breeding African rose beetles in his bathtub.
B: **You know ↗WHAT?!**

As highlighted in the examples above, we can use both focus (marked by capital letters) and intonation (marked by rising or falling arrows) as cues to tell these three constructions apart. In the constructions in (10) and (11), the Wh-word *what* needs to be unfocussed to derive the desired indefinite interpretation.³ This generalization seems to hold cross-linguistically (cf. Haida 2007; Roelofsen et al. 2019; Hengeveld et al. 2023). In contrast, in the echo questions presented in (12), the Wh-word has to be focussed. The latter two constructions can be further distinguished by their intonational contours. While the ellipsis constructions come with falling intonation, the echo questions showcase rising intonation.

³ Interestingly, the ellipsis constructions still seem to involve focus, but it falls on the verb, not the Wh word.

In addition to these contrasts, the three constructions also differ functionally. In the indefinite constructions in (10), the Wh-word *what* acts as an indefinite which seems to point towards the next sentence(s) uttered by the speaker. In (10a), for instance, the *what* in the *tell you what* construction appears to be co-referential with the following utterance *I have an idea*.

In the ellipsis constructions shown in (11), on the other hand, *what* functions as a Wh-complementizer that introduces an elided string. Usually, these constructions are preceded by a question from which we can easily and unambiguously recover the elided material. This is illustrated in (13) and (14).⁴

(13) What do you think **President Trump had to do with it?** I'll tell you **what** [~~President Trump had to with it~~]. Like, how about everything?

(14) NUDIE: Shit. What is this? Half rot-gut?
 HANK: What **are you talking about?**
 NUDIE: **You know what** [~~I am talking about~~]. This here's more booze than soda.

Last, in the echo question constructions in (12), the Wh-word *what* acts as an interrogative. Here, the speaker questions a surprising proposition by repeating it partially and leaving the Wh-word *what* in situ.

Of course, the three constructions in (10) to (12) also differ in other crucial regards, such as their degree of idiomaticity, their behaviour in the *something*-substitution test (see Section 1), and their role in discourse. However, due to spatial limitations, a proper discussion of these issues will have to be postponed.

⁴ While it might be tempting to argue that the indefinite constructions in (10) are also the result of ellipsis, such an account seems less convincing. One issue is that — if we assume that these constructions involve ellipsis — the elided material cannot be unambiguously recovered from a previous utterance, as exemplified in (ii).

- (ii) a. **You know what** [~~I just figured out~~] I just realized. That's my journal.
 b. **You know what** [~~I am so thrilled about~~] I just realized. That's my journal.
 c. **You know what** [~~my mother was so bummed out about~~] I just realized. That's my journal.

3 The *tell you what* construction

Having established how the indefinite *tell you what* construction differs from other similar looking utterances, this section will take a closer look at its form and function. I will argue that this construction is best described as a “lexically filled” idiom (cf. Fillmore et al. 1988) that speakers use to draw attention to an upcoming utterance. The examples in (15) to (18) illustrate its use.

- (15) Well, then we have something of an impasse. **I tell you what.** I’ll call the police — and what can I say? — if I’m wrong about the whole book-down-the-trousers scenario, I really apologize.
(*Notting Hill* [1999 film])
- (16) **I tell you what**, if they’d told me I could birth puppies from down there I might have gone for it maybe once, but doing it over and over just to get a human baby? I wasn’t doing it, No, no way!
(Sharon Fisher Corbett: *I Tell You What!*)
- (17) The only other food in the house is limes and Pop-Tarts. **Tell you what**, I’ll take you out to eat, how about that?
(Josephine Humphreys: *The Fireman’s Fair*)
- (18) Biff, first thing we gotta do when we get time is clip that big branch over the house. Afraid it’s gonna fall in a storm and hit the roof. **Tell you what.** We get a rope and sling her around, and then we climb up there with a couple of saws and take her down.
(Arthur Miller: *Death of a Salesman*)

While the examples above all represent the declarative version of the construction, it is worth noting that a hortative variant also exists, as shown in (19).

- (19) **Let me tell you what**, friends. With that kind of leader, it’s no wonder that Enron crashed and burned like no other corporation in American history. (MSNBC: “Scarborough” [2006-04-28])

However, for reasons of space, I will disregard this hortative variant for the rest of this paper and instead focus on the declarative realizations.

The indefinite *tell you what* construction, as shown above in (15) to (18), can be classified as a lexically filled idiom, as it allows little to no

variation in terms of its component parts. The subject slot needs always to be instantiated by the first-person singular pronoun *I* (which may be realized either overtly or covertly), and the verb slot can only be filled by the verb *tell*. While this verb usually surfaces as a simple present form, it may occasionally also occur with the *will* future or the *going-to* future, as highlighted by (20) and (21).

(20) So, you guys think I should have kissed her? Well, **I'll tell you what**. I'm gonna go kiss her. Right now.

(How I Met Your Mother: "Game Night")

(21) Well, **I'm going to tell you what**. You're going to go ahead and write me a dinner poem, and I'll belt that out after I choke this down.

(The Change Up [2011 film])

Furthermore, the indirect object needs to be realized by the second-person singular pronoun *you*, and the direct object can only be instantiated by the unfocussed Wh-word *what*. Any deviations from these tenets render the construction infelicitous, as shown in (22) to (27) below.⁵

(22) *Infelicity due to inappropriate subjects:*

- a. #**You** tell you(rself) what. You'll call the police.
- b. #**She** tells you what. She'll call the police.
- c. #**We** tell you what. We'll call the police.
- d. #**They** tell you what. They'll call the police.
- e. #**Briony** tells you what. She'll call the police.

(23) *Infelicity due to inappropriate verbs:*

- a. #**I say** (to) you what. I'll take you out to eat, how about that?
- b. #**I inform** you what. I'll take you out to eat, how about that?
- c. #**I propose** (to) you what. I'll take you out to eat, how about that?
- d. #**I suggest** (to) you what. I'll take you out to eat, how about that?

⁵ The attentive reader will have noticed that the *tell you what* construction meets the criteria for performative sentences, as proposed by Austin (1961). It requires a first-person singular subject, involves a performative verb (here: *tell*), and usually employs the simple present.

(24) *Infelicity due to inappropriate tense and aspect:*

- a. #I **told** you what. I'm gonna go kiss her.
- b. #I **have told** you what. I'm gonna go kiss her.
- c. #I **had told** you what. I'm gonna go kiss her.
- d. #I'm **telling** you what. I'm gonna go kiss her.

(25) *Infelicity due to inappropriate indirect objects:*

- a. #I tell **me/myself** what. I'm gonna go kiss her.
- b. #I tell **her** what. I'm gonna go kiss her.
- c. #I tell **them** what. I'm gonna go kiss her.
- d. #I tell **Lady Macbeth** what. I'm gonna go kiss her.
- e. ?I tell **you guys** what. I'm gonna go kiss her.

(26) *Infelicity due to inappropriate direct objects:*

- a. #I tell you **who**. Georgia.
- b. #I tell you **where**. London.
- c. #I tell you **when**. On Sunday.

(27) *Infelicity due to inappropriate focus:*

#I tell you **WHAT**. I'll call the police.

From a discourse perspective, the indefinite *tell you what* construction serves as an attention getting device. Thus, it cannot stand on its own but must be followed by another sentence in the imminent speech situation. In (28), for instance, the speaker uses the *tell you what* construction to draw the addressee's attention to the subsequent suggestion *How about you go back to sleep?*

(28) **I'll tell you what. How about you go back to sleep**, and then maybe Daddy will show up in your dream and then he can chase that monster away.

(The Possession of Michael King [2014 film])

Remaining silent or postponing the follow-up sentence until another time outside of the speech situation renders the construction infelicitous, as shown in (29) and (30).

(29) # **I tell you what.** (*silence*)

(30) # Tomorrow, **I will tell you what.**

But what is the nature of the following utterance? According to the *Oxford English Dictionary* (OED 2023), the *tell you what* construction is “[u]sed to introduce a suggestion or proposal” or an “observation or comment”. An impressionistic survey of a small sample of data from the *Corpus of Contemporary American English* (COCA) supports this description. Example (31), for instance, shows a case where *tell you what* introduces a suggestion, while (32) shows a case where it calls attention to an observation.

(31) **I tell you what.** I so rarely get a kindred spirit in here, my dear.
May I make you some tea?

(Brenda Carre: *Embrace of the Planets*)

(32) **I’ll tell you what. He is goood lookin’.**

(*Thelma & Louise* [1991 film])

In addition, the construction also seems to be able to introduce other speech acts, such as expressives, as in (33), or promises, as in (34).

(33) **I tell you what. Fuck your plan.** Lou. (Life [1999 film])

(34) **I’ll tell you what. I’ll be back in a little while.**

(*Buffy the Vampire Slayer*: “Never Kill a Boy on the First Date”)

To get a more comprehensive picture of what kinds of follow-up speech acts the *tell you what* construction is compatible with, a full-scale corpus study would be in order. Such a study, however, lies beyond the confines of this paper.

4 The *you know what* construction

Just like the *tell you what* construction, the indefinite *you know what* construction can also be described as a lexically filled idiom that draws attention to an upcoming utterance. The examples in (35) to (38) illustrate its use.

- (35) “**Do you know what?**” said Grieg. “While you have slept so peacefully, I have set my penetrating and illuminating intellect upon a formidable problem that still lies ahead of us, despite all of our preparations.” (Matthew Vierling: “Return to Zero”)
- (36) **You know what**, Spike? The more I get to know you, the more I wish I didn’t. (*Buffy the Vampire Slayer*: “Lie to Me”)
- (37) Hey, Lou, **you know what?** These raindrops. They got legs. (Stephen Schottenfeld: “Artie Gottlieb: Consulting Philosopher”)
- (38) Hybrid intelligence. HI. **You know what?** Screw that. Sounds like a marriage between a dolphin and a Toyota. (Robert Grossbach: *An Idea Whose Time Had Come*).

As highlighted by these examples, the construction comes in the shape of a highly idiomatic yes/no interrogative. While the auxiliary *do* may or may not be overtly encoded, the rest of the construction is essentially fixed. The subject is always instantiated by the second-person singular pronoun *you*, while the verb slot needs to be filled by the verb *know* in the simple present. Last, the object slot needs to be realized by the unfocussed Wh-word *what* to derive the desired indefinite interpretation. The use of other subjects, verbs, or objects results in infelicity, as shown in (39) to (43).

(39) *Infelicity due to inappropriate subjects:*

- a. #**I** know what? Screw that.
- b. #**He** knows what? Screw that.
- c. #**Ella** knows what? Screw that.
- d. #**We** know what? Screw that.
- f. ? **You guys** know what? Screw that.

(40) *Infelicity due to inappropriate verbs:*

- a. #Hey, Lou, you **are aware of** what? These raindrops. They got legs.
- b. #Hey, Lou, you **realize** what? These raindrops. They got legs.
- c. #Hey, Lou, you **perceive** what? These raindrops. They got legs.

- (41) *Infelicity due to inappropriate tenses and aspects:*
- a. # You **knew** what? These raindrops. They got legs.
 - b. # You **have known** what? These raindrops. They got legs.
 - c. # You **had known** what? These raindrops. They got legs.
 - d. # You **will know** what? These raindrops. They got legs.
 - e. # You **are going to know** what? These raindrops. They got legs.
- (42) *Infelicity due to inappropriate objects:*⁶
- a. # Hey, Lou, you know **who**? Anna.
 - b. # Hey, Lou, you know **when**? In January.
 - c. # Hey, Lou, you know **where**? Regensburg.
- (43) *Infelicity due to inappropriate focus:*
- # Hey, Lou, you know **WHAT**? These raindrops. They got legs.

From a discourse perspective, the *you know what* construction strongly resembles the *tell you what* construction in that it is also used to call attention to an upcoming utterance. In (44), for instance, the speaker employs the *you know what* construction to introduce the suggestion *Why don't you come out to Los Angeles and see for yourself what kind of a mother I am*.

- (44) “Hey, **you know what?**” Cee Cee said. “Yeah?” “**Why don't you come out to Los Angeles and see for yourself what kind of a mother I am.**”
(Iris Rainer Dart: *I'll Be There*)

Once again, the use of the construction would be infelicitous if the speaker did not provide a follow-up sentence but instead kept silent, as shown in (45).

- (45) # **You know what?** (*silence*)

The *OED* (2023) doesn't link the *you know what* construction to any follow-up speech act in particular, but describes it as being used “to emphasize or call special attention to what is said”. Indeed, a look at some corpus data from the *COCA* suggests that this construction can be

⁶ These utterances would be fine if we were talking about the ellipsis construction discussed in Section 2.

followed by a wide range of different speech acts, such as offers, as in (46), observations, as in (47), or bets, as in (48).

(46) Esther said, “**you know what?** I have an idea. **I could buy you a cappuccino**, in exchange for the cigarette. I mean, if you’re not busy.”
(Nino Ricci: *The Origin of Species*)

(47) Sure, she had her flaws, but **you know what? The girl had heart.**
(*The Wonder Years*: “Nemesis”)

(48) **You know what? I’ll prove it to you.** I’ll trade you Joey for Rachel and I’ll still win the game.
(*Friends*: “The One with the Football”)

Again, it would be intriguing to conduct a more thorough examination of what kinds of speech acts may or may not co-occur with this construction.

5 The history of the constructions

In the previous sections, I have shown that the *tell you what* and the *you know what* constructions play a special role in English, as proper Wh-indefinites otherwise do not seem to exist in the language. Naturally, this raises some questions concerning their origin. Have these two constructions been around for centuries and represent the last surviving remnants of a once productive quexistential system, or are they rather the product of a more recent innovation in English? To shed light on this matter, I conducted a diachronic corpus survey, examining data from several historical corpora covering the period from Old English to Modern English.⁷

Based on the corpus data, the *tell you what* construction has been in use at least since the Early Modern English period. More specifically, its first attested occurrence can be found in Abraham Hartwell’s translation of a Latin letter written around 1565, replicated below in (49).

⁷ In particular, I consulted the following resources: *The Dictionary of Old English Corpus (DOE)*, *The York-Toronto-Helsinki Parsed Corpus of Old English Prose (YCOE)*, the second edition of *The Penn-Helsinki Parsed Corpus of Middle English (PPCME2)*, *The Penn-Helsinki Parsed Corpus of Early Modern English (PPCEME)*, *The Corpus of Historical American English (COHA)*, *The Corpus of Contemporary American English (COCA)*, and *Google’s Ngram Viewer*.

- (49) As for Diuynitie, **I wyll tell you what**. it is so handled of .ii. men, in .ii. bookes, within these .ii. yeres, that better it had bene the gospel had neuer peped out.

(Abraham Hartwell: *A Sight of the Portugall Pearle* | ?1565)

In the following decades, the construction seemingly vanished again, until it re-emerged and mushroomed towards the end of the 16th century in the works of famous Elizabethan authors like William Shakespeare, Robert Greene, and Thomas Deloney. A selection of examples from this period is given in (50) to (52).

- (50) **Ile tell you what**, I thought my selfe as a proper fellow at wasters, as any in all our village, and yet when my wife begins to plaie clubbes trumpe with me, I am faine to sing:

(Robert Greene: *Selimus, Emperour of the Turkes (Part 1)* | 1594)

- (51) **I tell yee what**: Thursday is neere,
Lay hand on heart, aduise, bethink your selfe,
If you be mine, Ile giue you to my frend:

(William Shakespeare: *Romeo and Juliet* | 1597)

- (52) **I will tell thee what** (quoth Gillian) that man which needeth neither to flatter with his friends, nor borrow of his neighbours hath riches sufficient:

(Thomas Deloney: *The Pleasant Historie of Iacke of Newberie* | 1597)

Obviously, it is difficult to tell whether these examples represent an innovative use by the mentioned playwrights, or whether the sudden rise of the construction around 1600 simply reflects changing literary preferences and their effect on the composition of the consulted corpora. After all, during the Elizabethan era, plays became the most popular genre of literature. As plays are inherently dialogue heavy and consist almost exclusively of direct speech between two or more interlocutors, they offer an exceptionally fertile ground for the *tell you what* construction — perhaps more so than the genres that account for most of the textual material before then. This, however, remains pure speculation.

The *you know what* construction, on the other hand, appears to be much younger in comparison. The earliest attested instance of this

construction comes from a play written by Denman Thompson in 1885, as reproduced in (53).

- (53) RICKETY: Say, **do you know what?**
FRANK H.: No, what is it?
RICKETY: Well, I can climb a tree jest as good as a boy, — want to see me?
(Denman Thompson: *The Old Homestead* | 1885)

After that, the construction dropped off the radar for several decades, until it slowly started to gain momentum in the early 1920s and has increased in frequency ever since. Examples (54) to (56) show some of the early uses of the construction.

- (54) Wonderful! The water, dripping from you, must have looked like pearls. **Do you know what?** You're some sea goddess and you're only fooling us. (Harold MacGrath: *The Ragged Edge* | 1921)

- (55) **Do you know what?** He was a thief; he was stealing this auto.
(Percy Keese Fitzhugh: *Pee-Wee Harris on the Trail* | 1922)

- (56) MR. ZERO: Say, **do you know what?**
DAISY DIANA DOROTHEA DEVORE: What?
MR. ZERO: It makes me feel like dancin'.
(Elmer Leopold Rice: *The Adding Machine* | 1923)

While these observations suggest that the *you know what* construction might be a fairly recent innovation, it is worth mentioning that, several centuries earlier, a very similar construction already existed: the Middle English *wot ye what* construction. This construction is functionally identical to the Modern English *you know what* construction but differs in the verb it selects. More specifically, it does not involve the verb *know*, but the now obsolete Middle English verb *witen* — a cognate of Modern German *wissen* 'to know' and Modern Dutch *weten* 'to know'. Examples (57) and (58) illustrate its use in Middle English.

- (57) Ye be lyke the swynt catte That wolde haue fissh, but **wostow whatte?** He wold no thinge wete his clowes.
(Geoffrey Chaucer: *House of Fame* | 1380/1450)

- (58) In her presence we kneled down echon,
 Presentinge up our billes, and, **wot ye what**,
 Ful humb belly she took hem, by on and on;
 (Anonymous: *The Assembly of Ladies* | ca. 1400–1500)

This *wot ye what* construction survived into the Early Modern English period, as shown in (59) and (60), before it slowly began to vanish.

- (59) **Wot you what?** To day the Lords you talke of, are beheaded.
 (William Shakespeare: *Richard III* | 1623)
- (60) I found him at the market full of woe, crying a lost daughter, and telling all her tokens to the people; and **wot you what?** by all subscription in the world, it should be our new maid *Melvia*, one would little think it, therefore I was bold to tel him of her Mistriss.
 (Francis Beaumont & John Fletcher: *The Coxcomb* | 1647)

All things considered, the diachronic corpus survey thus provides solid evidence that idiomatic quexistential constructions, like the ones discussed in this paper, have been used by English speakers at least since the Middle English period. How productive the quexistential system really was at that point, however, remains an open question.

6 Conclusion

In this brief survey, I examined the indefinite *tell you what* and *you know what* constructions in English. Drawing on language data from several sources, I showed that both constructions can be classified as lexically filled idioms which speakers tend to use in discourse to draw attention to the subsequent utterance. A diachronic corpus study further suggests that constructions like these have been in use at least since the Middle English period, showing that English has had proper quexistentials for several centuries — at least in a small set of fixed expressions.

References

- Austin, John L. 1962. *How to Do Things with Words*. Oxford, UK: Clarendon Press.
- Fillmore, Charles, Paul Kay, & Mary Catherine O'Connor. 1988: Regularity and Idiomaticity in Grammatical Constructions: the Case of Let Alone. *Language* 64(3):501–538.

- Haida, Andreas. 2007. *The indefiniteness and focusing of wh-words*.
Doctoral dissertation, Humboldt University, Berlin, Germany.
- Hengeveld, Kees, Sabine Iatridou, & Floris Roelofsen. 2023.
Quexistentials and Focus. *Linguistic Inquiry* 54(3):571–624.
Oxford English Dictionary. 2023. URL: <https://www.oed.com/>.
- Roelofsen, Floris, Sabine Iatridou, & Kees Hengeveld. 2019.
Quexistentials and Focus. Presented at the University of British
Columbia and at Princeton.

Imperatives and prohibitives in Biblical Hebrew*

ELIZABETH RITTER
University of Calgary

1 Introduction

Directive sentence types are sometimes distinguished by polarity. IMPERATIVES are positive directives that request or command the addressee to undertake a particular action, as illustrated in (1a). PROHIBITIVES are negative directives that request or command the addressee to refrain from undertaking a particular action, as illustrated in (1b).

- (1) a. Hotze, please tell me about imperatives and prohibitives.
- b. Don't tell me you've never thought about this problem.

Imperatives have been widely studied both in the typological and the theoretical literature. The consensus is that (i) imperatives are universally attested; (ii) they typically permit or require a null subject whose discourse referent is the addressee; and (iii) in languages with so-called rich agreement, imperative verbs are typically not inflected for person (van der Auwera & Lejeune 2013b). Prohibitives are also known as negative imperatives, a term which would seem to imply that they are simply a subtype of imperative.

The primary goal of this squib is to compare the properties of imperatives and prohibitives in Biblical Hebrew, in order to determine the relationship between them. First, I show that imperatives and prohibitives have different verb forms. Then I show that this language has one type of imperative, but two types of prohibitives. This would seem to indicate that prohibitives are not simply negated imperatives. This is not a new idea. For example, Birjulin and Xrakovskij (2001:37) propose that imperative and prohibitive “paradigms should be viewed as independent, although semantically related entities.” In Section 3, I very briefly outline some of the questions raised by the facts of Biblical

* This squib was inspired by and written for my dear friend and colleague, Hotze Rullmann. Our conversations about syntax, semantics, and their interface are among my favourite memories of time well-spent doing linguistics. As I was writing about this little problem, I kept thinking how much I wished I could have talked it through with you. I hope one day I will.

Hebrew — and by comparable facts in other languages — that have not, to my knowledge, received a satisfactory answer in the existing literature.

2 The morphosyntax of Biblical Hebrew directives

In this section I describe the distinctive morphology of the verbs, negation markers and discourse particles that appear in Biblical Hebrew directives. I begin with verbs, which are typically inflected for person, number, and gender. This is the case for both the perfect and imperfect verb forms.¹ However, imperative verbs are inflected for number and gender only. They are identical to the imperfect second person forms, except that they lack the person prefix (*ti-*) (van der Merwe et al. 1999). See, for example, the imperative and imperfect forms of the verb ‘write’ in Table 1 below:²

Table 1: Imperfect and imperative forms of verb ‘write’³

	Imperfect	Imperative
2MSG	ti-kətov	kətov
2FSG	ti-ktəvi	kitvi
2MPL	ti-ktəvu	kitvu
2FPL	ti-kətovna	kətovna

(adapted from van der Merwe et al. 1999:70–71)

Like many languages, Biblical Hebrew does not use imperative verb forms in its prohibitive sentences (van der Auwera & Lejeune 2013a). Moreover, the form of the verb in a prohibitive sentence depends on whether or not the prohibition is specific to the discourse situation and

¹ Perfect and imperfect are labels for verb paradigms and are often referred to as tenses in the Biblical Hebrew literature, though the choice between them typically depends on aspectual considerations (cf. Pratico and van Pelt 2007:130). The perfect is used for states of being and for completed events. The former is typically translated into English with present tense verbs, and the latter with past tense verbs. In contrast, the imperfect is used for incomplete events.

² The following abbreviations are used in this paper: 1/2/3: first/second/third person; DEF definite; DEM demonstrative; EMPH emphatic; F feminine; IMPF imperfect; IMPTV imperative; INF infinitive; JUSS jussive; M masculine; NEG negation; O object; OM object marker; PART participle; PERF perfect; PL plural; SG singular.

³ The vowel pattern alternations that distinguish the imperfect and imperative in the feminine singular and masculine plural verb forms are entirely predictable, and hence, are not considered relevant for the characterization of the similarities and differences between these two sets of verb forms. Specifically, the [i] results from the fusion of two audible instances of [ə] (van der Merwe et al. 1999:71).

the intended addressee. Henceforth, I will refer to these as specific and nonspecific prohibitives, respectively. Specific prohibitives require the jussive form of the verb, whereas nonspecific prohibitives require the imperfect form of the verb.⁴ (In Biblical Hebrew, the jussive is identical to the imperfect form for most verbs, including the verb ‘write’ as in Table 1, and otherwise is a short form of the imperfect derived by apocoping the final vowel.)

Turning next to negation, Biblical Hebrew has two markers of negation, *’al* and *lo*, and their distribution is as follows: *’al* only occurs in specific prohibitives, *lo* occurs elsewhere — it is used both for clausal and constituent negation. The examples in (2) and (3) illustrate the use of *’al*. Note that in both examples the speaker is instructing the current addressee to refrain from doing something. The examples in (4), which are from the ten commandments, illustrate the use of *lo* in non-specific prohibitives; they apply to anyone at any time. Similarly, the example in (5), which also contains *lo*, explicitly states that the prohibition applies to everyone, everywhere and at all times. The examples in (6) demonstrate that *lo* is the default clausal negation marker, as they illustrate the uses of this particle in declaratives with perfect and imperfect verb forms.

- (2) w^a-‘āšārâ ’ānāšîm niməṣə’ û-b-ām
 and-ten people found.3PL and-among-3PL
 wa-yyō’ mərû ’el-yišəmā’el ’al-təmit-ēnû
 and-say.IMPF.3PL to-Ishmael NEG-kill.JUSS.2SG-1PL.O
 ‘But ten men were found among them who said to Ishmael, “Do not kill us, ...” [Jeremiah 41:8]

- (3) ’aḥar ha-dəbārîm hā-’ēlleh hāyâ dəḅar-yəhwâ
 after DEF-things DEM-PL be.IMPF.3SG word-God
 ’el-’abərām ba-mm^aḥāzeh lē’mōr ’al-tîrâ’
 to-Abram in-vision say.INF NEG-fear.JUSS.2MSG
 ’abərām ...
 Abram ...
 ‘After these things the word of the Lord came unto Abram in a vision, saying, “Fear not, Abram ...” [Genesis 15:1]

⁴ Different authors use different terms for these two types of prohibitives. For example, Van der Merwe et al. (1999) refer to specific and nonspecific prohibitives as direct and indirect prohibitives, respectively. Waltke and O’Connor (1990) characterize specific prohibitives as indicating urgency in contrast to nonspecific prohibitives, which they characterize as indicating legislation.

- (4) **lō'** tirəṣāḥ: **lō'** tinə'āf:
 NEG kill.IMPF.2MSG NEG commit.adultery.IMPF.2MSG
lō' tiḡənōḇ:
 NEG steal.IMPF.2SG
 'Thou shalt not kill. Thou shall not commit adultery. Thou shalt not steal.'
 [Exodus 20:13–14]
- (5) wə-leḥem wə-qālī wə-ḵarəmel **lō'**
 and-bread and-roasted.grain and-fresh.grain NEG
 tō'ḵəlū 'aḍ-'ešem ha-yywōm hazzeh 'aḍ
 eat.IMPF.2MPL until-EMPH DEF-day DEM until
 ḥāḇī'ākem 'et-qārəban 'əlōhē-ḵem ḥuqqat
 bring.PERF.2MPL OM-sacrifice god-2MPL law
 'wōlām lə-dōrōtē-ḵem bə-ḵōl
 eternal to-generations-2MPL in-all
 mōšəḇōtē-ḵem:
 dwelling.places-2MPL
 'You shall not eat bread or [flour made from] roasted grain or fresh grain, until this very day, until you bring your God's sacrifice. [This is] an eternal statute throughout your generations in all your dwelling places.'
 [Leviticus 23:14]
- (6) a. wa-'āḍabərā bə-'ēḍōtēy-ḵā neḡeḍ
 and-speak.IMPF.1SG of-testimonies-2MSG against
 mālāḵīm wə-**lō'** 'ēḇwōš:
 kings and-NEG be.ashamed.IMPF.1SG
 'And I shall speak of Your testimonies in the presence of kings, and I shall not be ashamed.'
 [Psalms 119:46]
- b. mi-twōrāt-əḵā **lō'** nāṭīṭī:
 from-torah-2MSG NEG turn.PERF.1SG
 'I did not turn away from Your Torah.'
 [Psalms 119:50]

Significantly, while there are two types of prohibitives, there is only one type of imperative. This can be seen by comparing the examples in (7) and (8). The former is a specific command issued by Abraham to his servant, and the latter is a positive commandment, that is, a non-specific command. Both contain an imperative verb form.

- (7) wayyō'mer ' aḇərāām 'el-'aḇəd-wō zəqan
 and.say.IMPF.3MSG Abraham to-servant-3MSG old
 bêt-wō ha-mmōšēl bə-kāl-'āšer-lwō
 house-3MSG DEF-rule.PART over-all-that-3MSG
 śīm-nā' yād-əkā t^aḥat̄ yərēk-î:
 put.IMPTV.MSG-NA hand-2MSG under thigh-1SG
 'So Abraham said to the oldest servant of his house, who ruled
 over all that he had, "Please, put your hand under my thigh."
 [Genesis 24:2]
- (8) kabēd 'et-'ābî-kā wə-'et-'imm-ekā
 honor.IMPTV.MSG OM-father-2MSG and-OM-mother-2MG
 'Honour thy father and thy mother.' [Exodus 20:12]

One final element that is only seen in specific directives is the particle *nā*. This particle is variously analysed as a particle of entreaty, translated as 'please' or 'I pray' or 'I beg you' (Gesenius & Kautzsch 1909; Kaufman 1991), as a logical particle that is best left untranslated (Lambdin 1971; Waltke & O'Connor 1990) or as a propositive particle that signals speaker intention to pursue a particular course of action (Shulman 1999; Christiansen 2009). When it occurs in imperatives, *nā* is suffixed to the imperative verb, as illustrated in (7) above, and when it occurs in specific prohibitives, it is suffixed to the negative particle 'al, as illustrated in (9).⁵

- (9) wa-yyō'mer 'al-nā t^a'āzōḇ 'ōṭānū ...
 and-say.IMPF.3MSG NEG-NA leave.JUSS.2MS OM.1PL
 'and he said "Please do not leave us".' [Numbers 10:31]

Significantly, *nā* never occurs in nonspecific prohibitives. In other words, there are no examples of *nā-lo* in the Hebrew Bible.

The following table summarizes the properties of imperatives, specific prohibitives and nonspecific prohibitives described above:

⁵ The particle *na* can also occur on the complementizer *im* 'if', as illustrated in (i):

- (i) wayyō'mar 'ādōnāy 'im-nā' māšā'-tî ḥēn bə-'ēney-kā 'al-nā'
 and.said my.lord if-NA found-1SG favour in-sight-2SG NEG-NA
 t^a'āḇōr mē'al 'aḇəde-kā:
 pass.away from servant-2SG
 '... and said, My lord, **if now** I have found favor in thy sight, pass not away, I pray
 thee, from thy servant' [Genesis 18:3]

Table 2: Properties of Biblical Hebrew Directives

	Imperatives	Specific Prohibitives	Nonspecific Prohibitive
Verb form	imperative	jussive	imperfect
Verb agreement	number, gender	2nd person, number, gender	2nd person, number, gender
Negation marker	n/a	'al (prohibitives only)	lo (default clausal negator)
Entreaty particle	nā	nā	NONE

In the next section I identify some of the questions raised by this array of properties regarding the morphosyntax of these three types of directives, and the similarities and differences in their semantic interpretation.

3 Questions (for Hotze) about imperatives and prohibitives

The facts described in the last section raise a number of questions. First, why do Biblical Hebrew imperatives and prohibitives require different verb forms? Is this simply because when negation is present it blocks some kind of syntactic feature checking or movement operation that is obligatory in imperative clauses? Various researchers have suggested explanations along these lines hypothesizing that imperative verbs check a feature in a higher functional head, such as Mood or C or Force (e.g., Rivero 1994; Rivero & Terzi 1995; Zanuttini 1997).

Note, however, that — as is the case in many languages — Biblical Hebrew imperative verbs are not just different from verbs in other paradigms, they are *defective* in the sense that they lack person features. What is the significance of this defect? Does the absence of person in imperative verbs tell us that they are inflected for imperative force or imperative mood, but not tense/aspect? If so, should we interpret the lack of defective verbs in prohibitives as an indication that these sentences are not in fact imperative? And if so, are they inflected for tense/aspect, rather than (imperative) force or mood?

Imperatives are commands, requests, suggestions, etc. to *do* something. However, prohibitives are commands, requests, suggestions, etc. to *refrain* from doing something. This would seem to be the crux of

the difference between imperatives and prohibitives. Does this mean that they constitute different illocutionary acts? For example, Portner (2004) analyses imperatives as instructions to add an item to the addressee's to-do list, but an analysis along these lines cannot be straightforwardly extended to prohibitives: We all have to-do lists that are longer than they should be, and they often look like a list of imperatives, for example, *buy groceries, finish this squib*. But who keeps a to-don't list, and what would that look like?

Another set of questions is raised by the fact that there are two types of prohibitives. My intuition is that what I am calling specific prohibitives are uttered when the speaker believes that the addressee would otherwise do whatever it is they are telling them not to do in the (near) future. Nonspecific prohibitives, on the other hand, require no such belief on the part of the speaker. In other words, they are not restricted to the current addressee or the current discourse situation. As has often been noted, they typically express rules or laws to be followed by everyone in every relevant situation. This can be illustrated with the English examples below:

- (10) a. Don't park here! (This means YOU.)
 b. No parking.

The example in (10a) is a specific prohibitive. It is something that a curmudgeon might post on his back fence to keep his next-door neighbours from parking there — in other words, it is an instruction to be interpreted as immediate and personal. The example in (10b), on the other hand, is a nonspecific prohibitive. It could be posted by a municipality or other institution to indicate that a particular area was not available for parking by anyone at any time. I suspect that the contrast in (10), and similar pairs in Biblical Hebrew provided above, differ in that the specific prohibitive requires a representation of the current addressee and the current discourse situation, while these elements are missing in the representation of non-specific prohibitives. See Ritter and Wiltschko (2019) for a similar treatment of personal and impersonal *you*.

Additional support for a distinction along these lines comes from the observation that the particle *nā* is unavailable for non-specific prohibitives. While its precise semantic contribution to the sentences in which it appears is subject to debate, it frequently expresses something about the speaker's intentions or some aspect of their relationship to the current addressee, much like English *please*. Woods (2021) proposes that *please* alternatively marks a sentence as a request or reinforces the

speaker's and addressee's respective roles in the request. In either case, she argues that it is part of sentence structure, and as such is syntactically represented as a head in what Miyagawa (2022) refers to as the syntactic treetop — the topmost layer of syntactic structure, whose function is to represent aspects of the speech act (Speas & Tenny 2003), or conversational interaction (Wiltschko 2021), including the speaker, the addressee, and the illocutionary force. If an explanation along these lines is correct, it begs the question as to whether there are also (perhaps more subtle) differences between specific and non-specific imperatives. Think about that. But since this is for you, Hotze, perhaps I should say, think about that, please, won't you, Hotze?

References

- Birjulin, Leonid & Victor S. Xrakovskij. 2001. Imperative sentences: Theoretical problems. In Victor S. Xrakovskij (ed.), *Typology of Imperative Constructions*. Munich, Germany: Lincom, 3–50.
- Christiansen, Bent. 2009. A linguistic analysis of the Biblical Hebrew Particle *nā'*: A test case. *Vetus Testamentum* 59(3):379–393.
- Gesenius, Wilhelm & Emil Kautzsch. 1909. *Gesenius' Hebrew Grammar*. 2nd English edition, revised in accordance with the 28th German edition (1909) by A. E. Cowley. Oxford, UK: Oxford University Clarendon Press.
- Kaufman, Stephen A. 1991. An emphatic plea for *please*. *Maarav* 7(1): 195–198.
- Lambdin, Thomas O. 1971. *Introduction to Biblical Hebrew*. New York, NY: Scribner.
- Miyagawa, Shigeru. 2022. *Syntax in the Treetops*. Cambridge, MA: MIT Press.
- Pratico, Gary D. & Miles V. Van Pelt. 2007. *Basics of Biblical Hebrew*. Grand Rapids, MI: Zonderfan.
- Portner, Paul. 2004. The semantics of imperatives within a theory of clause types. In Robert B. Young (ed.), *Proceedings of SALT* 14:235–252.
- Ritter, Elizabeth & Martina Wiltschko. 2019. Nominal speech act structure: Evidence from the structural deficiency of impersonal pronouns. *Canadian Journal of Linguistics* 64(4):709–729.

- Rivero, María-Luisa, 1994. Negation, imperatives and Wackernagel effects. *Rivista di Linguistica* 6:39–66.
- Rivero, María Luisa & Arhonto Terzi. 1995. Imperatives, V-movement and logical mood. *Journal of Linguistics* 31(2):301–332.
- Shulman, Ahouva. 1999. The Particle ׀ in Biblical Hebrew Prose. *Hebrew Studies* 40(1):57–82.
- Speas, Margaret & Carol Tenny. 2003. Configurational properties of point of view roles. In Anna Maria Di Sciullo (ed.), *Asymmetry in Grammar*. Amsterdam, Netherlands: John Benjamins, 315–344.
- Waltke, Bruce K. & Michael O'Connor. 1990. *An Introduction to Biblical Hebrew Syntax*. Winona Lake, IN: Eisenbrauns.
- van der Auwera, Johan & Ludo Lejeune (with Valentin Goussev). 2013a. The Prohibitive. In Matthew S. Dryer & Martin Haspelmath (eds.), *WALS Online* (v2020.3) [Data set]. Zenodo. URL: <http://wals.info/chapter/71>. Accessed on 2023-08-09.
- van der Auwera, Johan & Ludo Lejeune (with Umarani Pappuswamy, Valentin Goussev). 2013b. The Morphological Imperative. In Matthew S. Dryer & Martin Haspelmath (eds.), *WALS Online* (v2020.3) [Data set]. Zenodo. URL: <http://wals.info/chapter/70>. Accessed on 2023-08-09.
- van der Merwe, Christo H. J., Jackie A. Naudé, & Jan H. Kroeze. 1999. *A Biblical Hebrew Reference Grammar*. Sheffield, UK: Sheffield Academic Press.
- Wiltschko, Martina. 2021. *The grammar of Interactional Language*. Cambridge, UK: Cambridge University Press.
- Woods, Rebecca. 2021. Towards a model of the syntax–discourse interface: a syntactic analysis of please. *English Language & Linguistics* 25(1):121–153.
- Zanuttini, Raffaella. 1997. *Negation and Clausal Structure: A Comparative Study of Romance Languages*. Oxford, UK: Oxford University Press.

Ktunaxa Ranges*

STARR SANDOVAL

University of British Columbia

1 Introduction

This paper discusses the expression of ranges in Ktunaxa.^{1,2} I review how phrases that express upper and lower bounds can be used in spatial, temporal, and degree domains. I use the empirical and theoretical frameworks for ranges presented in Gobeski and Morzycki (2022) and expanded upon

* This paper is for Hotze — whom I have been so lucky to know, learn from, and work with for the past three years. I will always be grateful for how welcoming and empathetic he was towards me when I was a nervous first year starting graduate school online, and his intellectual perspective has been invaluable as I have worked with him to develop my research. I have really loved our meetings and sharing so many linguistic/semantics spaces with him — whether we are talking about as-phrases, relational adjectives, the free indirect discourse of James Joyce, or stroopwafels. Thank you to Hotze for all of the academic and moral support he has provided me and so many others — this department and this field have benefited immensely from his contributions.

¹ Ktunaxa is a language isolate spoken in parts of interior British Columbia, and northern Montana, Idaho, and Washington. According to the 2022 FPCC Language Status Report, there are 18 fluent speakers in British Columbia, which classifies Ktunaxa as severely endangered (Gessner et al. 2022).

² The following data comes from my own primary fieldwork with Ktunaxa speaker Violet Birdstone. The working language for conducting interviews was English, and the primary elicitation strategies used were translation tasks and acceptability judgment tasks (Bochnak and Matthewson 2015). A sentence presented without a diacritic signals the sentence is semantically well-formed in a given context. # signals that a sentence was judged as semantically ill-formed in a given context. A check mark, ✓, next to a context signals that the sentence is compatible with the context in the example and an x mark, ✗, indicates that the sentence is in compatible with a given context. The following shorthands are used in morpheme glosses: 1 = 1st person, 3 = 3rd person, IND = indicative, OBV = obviative, PRVB = preverb, SUBJ = subject, COMP = complementizer, COP = copula, DEM = demonstrative, DIST = distal demonstrative, GRP = group, INDEF.AMT = indefinite amount, NEG = negative, PASS = passive, PST = past, RFLX = reflexive. Characters in Ktunaxa orthography not conforming to their typical IPA values are as follows: † = [t̥], ɕ = [t̥s].

by Sandoval (2023). I discuss one productive range construction in Ktunaxa (though there may be others). It appears to be derived from a locative phrase signaling a start- and an endpoint (similar to how English *from...to* functions). In the subsequent sections, I discuss background on range expressions, the ingredients used for ranges in Ktunaxa, and their application to spatial, temporal and degree ranges.

2 Degree and temporal ranges in English

Gobeski and Morzycki (2022) classify ranges as a subkind of *composite measure phrase*—measure phrases with multiple measure phrases as sub-constituents. Ranges have two arguments—one that sets a lower bound and one that sets an upper bound. Gobeski and Morzycki (2022) identify three distinct English expressions that refer to ranges: *from...to* (1a), *between...and* (1b), and *through* (1c). All of these phrases include two degrees (or names that map to an ordered list) that set an upper and lower bound of a range.

- (1) a. This volume spans **from** Lincoln **to** Taft. (*from d to d'*)
- b. This volume spans **between** Lincoln **and** Taft. (*between d and d'*)
- c. This volume spans Lincoln **through** Taft. (*d through d'*)

Gobeski and Morzycki also identify three different kinds of range readings that relate to the value of the degree under discussion. Under the *singleton punctual* reading, the range expression references a span that *one value* falls into. In (2a), Floyd's age is naturally a single, unique value, and the range expression *between four and six* references a scale that the single value falls on. Meanwhile, under the *set punctual* reading, the range expression references a scale that a set of multiple values fall onto. In (2b), there are multiple children, so there are multiple ages that fall into the range. Lastly, there is the *interval* reading, under which the discussed value spans the *entire* distance that the range expression denotes. In (2c), degrees at which the chemical freezes span the entirety of the range named.

- (2) a. Floyd is between four and six years old. (*singleton punctual*)
age 4 · · · age 6
- b. The children are between four and six years old. (*set punctual*)
age 4 · · · · age 6
- c. The chemical freezes between -15° and 0°. (*interval*)
-15° ————— 0°

Gobeski and Morzycki identify *from...to* and *between...and* range expressions as compatible with all three readings.³ Meanwhile *through* only allows an interval reading. Table 1 summarizes their generalizations.

	singleton punctual	set punctual	interval
from d to d'	✓	✓	✓
between d and d'	✓	✓	✓
d through d'	✗	✗	✓

Table 1: Range interpretations available with range expressions in the degree domain

Sandoval (2023) additionally discusses ranges in the temporal domain, which behave similarly to degree range constructions, though *from...to* fully resists a *singleton punctual* reading, shown in (3a). Meanwhile it improves under a *set punctual* reading and especially an *interval* reading, shown in (3b) and (3c). This data is summarized in Table 2.

- (3) a. # Clyde’s birthday is from Monday to Friday. (*singleton punctual*)
- b. The guests arrived from Monday to Friday. (*set punctual*)
- c. The event lasted from Monday to Friday. (*interval*)

³ There are contexts in which *from...to* resists a *singleton punctual* interpretation. For more discussion of this constraint, see Geurts and Nouwen (2007); Gobeski and Morzycki (2022); Kennedy (2015); Sandoval (2023).

	singleton punctual	set punctual	interval
from <i>d</i> to <i>d'</i>	✗	✓	✓
between <i>d</i> and <i>d'</i>	✓	✓	✓
<i>d</i> through <i>d'</i>	✗	✗	✓

Table 2: Range interpretations available with range expressions in the temporal domain

3 Morphological ingredients for Ktunaxa ranges in the spatial domain

Ktunaxa can express the start- and endpoints of a spatial path using *ʔi-s-ϕ* to express the beginning region of the path (i.e. ‘from’) and *ϕ qu* to express the destination of the path (i.e. ‘to’). An example sentence using this construction is shown in (4).

- (4) Mu ʕinaxi ʔisϕ Vancouver ϕ qu Toronto.
 Ma-hu ʕinax-i ʔi-s-ϕ Vancouver ϕ qu Toronto.
 PST-1.SUBJ go-IND DET-OBV-and Vancouver and DIST Toronto.
 ‘I went **from** Vancouver **to** Toronto.’

On its own, *ʔis* is a determiner marked with obviation. In (5), it modifies the NP *kapis*, ‘coffee’.

- (5) ʔapit ʔikuʔni ʔis kapis Mali
 ʔap-it ʔikuʔ-ni ʔi-s kapi-s Mali
 all-PRVB drink-IND DET-OBV coffee-OBV Mary
 ‘Mary drank all **the** coffee.’

ϕ on its own functions as a conjunction. In (6) it conjoins *hanuhusni* ‘red’ and *hamakϕiʔni* ‘yellow’.

- (6) Niʔi ʔaqukwuk hanuhusni ϕ hamakϕiʔni.
 Niʔi ʔaqukwuk hanuhus-ni ϕ hamakϕiʔni
 DET box red-IND and yellow-IND
 ‘The box is red **and** yellow.’

ʔis and *ϕ* co-occur as *ʔisϕ* to mark a standard of comparison in degree constructions Bertrand (2021). This is shown in (7).

- (7) La-t wuqatitqakni **?isç** ?amlus
 La-t wuqatitqak-ni ?i-s-ç ?amlu-s
 La-t tall-IND DET-OBV-and ?amlu-OBV
 ‘La-t is taller **than** ?amlu’

Qus functions as a distal demonstrative marked by obviation. In (8), it picks out the (location of the) swans and indicates they are relatively far from the speaker.

- (8) ?inu, **qus** ?ini hakisukki ?u?u
 ?inu, qu-s ?i-ni hakisukq-i ?u?u
 look, DIST.DEM-OBV COP-IND swim.by-IND swan
 ‘Look, **over there/those** are swans swimming by.’

Together these morphological ingredients can mark the start- and end-point of a region. A final relevant point is that all of these morphemes are obligatory, as shown in (9).

- (9) Mu çinaxi ***(?isç)** Vancouver ***(ç) *(qu)**
 Ma-hu ç'inax-i ?i-s-ç Vancouver ç qu
 PST-1.SUBJ go-IND DET-OBV-and Vancouver and DIST.DEM
 Toronto.
 Toronto.
 Toronto.
 ‘I went **from** Vancouver **to** Toronto.’

4 Degree ranges

Degree ranges are formed similarly to how spatial paths are expressed. *?isç* is still used to mark the start point/lower bound. However, only *ç*, not *ç qus*, is used to mark the upper bound. Using the distal demonstrative in these cases is semantically anomalous. Meanwhile *qus* is obligatory when describing spatial paths. The sentence in (10) shows this fact and additionally demonstrates this structure can be used with a *singleton punctual* reading. Furthermore, it is felicitous in both a context where the speaker does (Context 1) and does not (Context 2) know the exact number.

- (10) **Context 1:** *You know how many apples La-t ate, and you want your sister to guess. You give her a hint that it was between four and six.*
Context 2: *You don't remember how many apples La-t ate exactly, but it was somewhere between four and six.*

ʔikni	ʔisɕ	xaɕas	ɕ	(*qus)	ʔinmisa-s
ʔik-ni	ʔi-s-ɕ	xaɕa-s	ɕ	qu-s	ʔinmisa-s
eat-IND	DET-OBV-and	four-OBV	and	DIST.DEM-OBV	six-OBV
	kanuhusnanas	La-t.			
	kanuhusnana-s	La-t			
	apple-OBV	La-t			

‘La-t ate **from** four **to** six apples.’

This structure also allows for a *set punctual* reading. In (11) there are multiple values within the range, as there are multiple children who ate apples.

- (11) *Context: The children in your class all ate apples. Some had four, some had five, some had six.*

ʔikni	ʔisɕ	xaɕas	ɕ	(*qus)	ʔinmisa-s
ʔik-ni	ʔi-s-ɕ	xaɕa-s	ɕ	qu-s	ʔinmisa-s
eat-IND	DET-OBV-and	four-OBV	and	DIST.DEM-OBV	six-OBV
	kanuhusnanas	ni		ʔkamnintik.	
	kanuhusnana-s	ni		ʔkam-nintik.	
	apple-OBV	DETchild-GRP			

‘The children ate **from** four **to** six apples.’

5 Temporal ranges

To express a temporal range in Ktunaxa, the same construction is used as the spatial distance construction used in the previous section. *ʔisɕ* marks the start of the range and *ɕ qus* marks the end of the range. However, in temporal range constructions, the distal demonstrative *qus* is optional, whereas with the spatial examples, it is obligatory, and with the degree examples, it is impossible. (12) provides an example of an interval reading of a temporal range. My consultant commented that the sentence sounded well-formed with or without *qus* and judged both versions to mean the same thing.

- (12) *Context: La·t wanted to be healthier, so he did not eat sweets for the first three months of the year.*

La·t qa ʔikni kquqçils ʔikits ʔisç
 La·t qa ʔik-ni k-quqçil-s ʔik-it-s ʔi-s-ç
 La·t NEG eat-IND COMP-sweet-OBV eat-PASS-OBV DET-OBV-and
 kmitxaltitnams ç (qus) hikuks
 kmitxaltitnam-s ç qu-s hikuq-s
 January-OBV and DIST.DEM-OBV March-OBV
 ‘La·t didn’t eat sweets **from** January **to** March.’

Like *from...to* in English, temporal ranges in Ktunaxa resist singleton punctual readings.

- (13) *Context: You’re a detective and you find La·t dead. You don’t know when he died, but he was away Monday to Friday so it must have been sometime then.*

#ʔupni ʔisç kłaʔukinmiyits ç
 ʔup-ni ʔi-s-ç k-ła-ʔuki=n-miyit-s ç
 die-IND DEM-OBV-and COMP-again-one-day-OBV and
 (qus) kłayi·kunmiyits La·t
 qu-s k-ła-yi·ku=n-miyit-s La·t
 DIST.DEM-OBV COMP-again-five-day-OBV La·t
 Intended ‘La·t died between Monday and Friday.’

Consultant comment: This would mean he was dying the whole time.

My consultant volunteered a repair for (13) by inserting *ʔaqanmiyits* ‘some day’ into the structure. This also makes a *from...to* structure more felicitous in English according to some speakers (15) when it would otherwise be fully ill-formed.

- (14) *Context: You're a detective and you find La·t dead. You don't know when he died, but he was away Monday to Friday so it must have been sometime then.*

ʔupni ʔaqanmiyits ʔisϕ
 ʔup-ni ʔaqan-miyit-s ʔi-s-ϕ
 die-IND INDEF.AMT-day-OBV DEM-OBV-and
 kʷaʔukinmiyits ϕ (qus)
 k-ʷa-ʔuki=n-miyit-s ϕ qu-s
 COMP-again-one-day-OBV and DIST.DEM-OBV
 kʷayi·kunmiyits La·t
 k-ʷa-yi·ku=n-miyit-s La·t
 COMP-again-five-day-OBV La·t
 ‘La·t died sometime between Monday and Friday.’

- (15) Floyd died #(sometime) from Monday to Friday.

6 Taking stock

To summarize, Ktunaxa range expressions (and their related locative, region-referencing expressions) all require *ʔisϕ* to introduce the lower bound and *ϕ* to mark the upper bound. However, the inclusion of *qus* after *ϕ* in the upper bound is a point of variation. For spatial distances it is obligatory, for temporal ranges it is optional, and for degree ranges it is prohibited. These generalizations are summarized in Table 3.

	Obligatory	Optional	Prohibited
Spatial	✓		
Temporal		✓	
Degree			✓

Table 3: Distribution of *qus* across range expressions in Ktunaxa

Additionally, spatial and degree range expressions allow singleton punctual expressions, while temporal range expressions restrict them. The behavior of the spatial data contrasts with *from...to* in English, which does not allow a *singleton punctual* reading, shown in (16).

- (16) # The museum is from Boston to Cleveland.

This data is summarized in Table 4.

	from...to	ʔisɕ qus
degree	✓	✓
temporal	✗	✗
spatial	✗	✓

Table 4: Allowance of singleton punctual readings

Degree range expressions in Ktunaxa marked by *ʔisɕ...ɕ* liberally allow a *singleton punctual* interpretation, unlike English. In English, the felicity of *singleton punctual* readings of *from...to* degree range expressions is variable and there are constraints related to ignorance and potentially unknown other factors (Gobeski and Morzycki 2022). Whether temporal range expressions marked by *ʔisɕ...ɕ* (*qus*) allow *singleton punctual* readings without *ʔaqanmiyit* (the modifier ‘some day’) requires future research.

Ktunaxa is also different from English in that there appears to only be one method for forming range constructions. While English has *from...to*, *between...and*, and *through*, Ktunaxa only appears to have the equivalent of English *from...to*. This may be because only *from...to* can be expressed with free morphemes. For example, when I have tried to elicit *between...and* in a spatial context, my consultant used the verb *yankinmit*, ‘to separate’.

(17) *Context: You are seated between ʔamlu and La·t.*

Hun yankinmitni k sankamik ʔamlu ɕ La·t
 Hun yankinmit-ni k sankamik ʔamlu ɕ La·t
 I.SUBJ separate-IND COMP stand.RFLX ʔamlu and La·t
 ‘I separated ʔamlu and La·t with myself.’

A final cross-linguistic parallel is that in English, range expressions can be used somewhat figuratively to express that there exist a wide variety of options or events that are not necessarily in a linear order. Some examples are shown in (18).

- (18) a. We have flavors from blueberry to pistachio.
 b. Floyd’s emotions ranged from elated to angry.
 c. From breaking my leg to seeing a double rainbow, this has

certainly been an eventful trip.

These examples may mirror non-numeral uses of *at least/most* as discussed by Geurts and Nouwen (2007). An example is shown in (19).

(19) The trip was awful, but **at least** the weather was nice.

It appears that at least with range constructions, Ktunaxa also allows this figurative expression as shown in (20). There is no standardized scale with edges marked by blue jays and bears. Rather, as these two animals are distinguished from one another, they demonstrate that Vancouver hosts a variety of wildlife.

(20) *Context: Vancouver has a lot of wildlife. You want to express that it has a lot of different types of animals.*

Haqaʔni	kt'uq̓qamnas	ʔis̓	ququskis	ǰ
Haqaʔ-ni	kt'uq̓qamna-s	ʔi-s-ǰ	ququksi-s	ǰ
exist-IND	animal-OBV	DET-OBV-and	bluejay-OBV	and
nupqus	Vancouver.			
nupqu-s	Vancouver.			
bear-OBV	Vancouver			

‘Vancouver has animals from blue jays to bears.’

In summary, Ktunaxa range expressions show striking similarities to English with respect to spatial metaphor and available interpretations, but also prompt future questions about the use of the distal demonstrative *qus*, strategies for expressing locations, and the use of *ʔaqanmiyit* (‘some day’) modification. This data also prompts larger cross-linguistic questions about ranges—how frequently do degree and temporal ranges map to locative prepositions, and what similarities are there between spatial expressions used and the range readings they allow?

References

- Bertrand, Anne. 2021. Investigating comparatives in Ktunaxa: Fieldwork and crosslinguistic perspectives. UBC colloquium, April 9th.
- Bochnak, M. Ryan, and Lisa Matthewson. 2015. *Methodologies in semantic fieldwork*. Oxford University Press, New York.
- Gessner, Suzanne, Tracey Herbert, and Aliana Parker. 2022. The report on the status of B.C. First Nations Languages 2022.

- Geurts, Bart, and Rick Nouwen. 2007. ‘At least’ et al.: the semantics of scalar modifiers. *Language* 83:533–559.
- Gobeski, Adam, and Marcin Morzycki. 2022. Ranges: composite measure phrases, modified numerals, and choice functions. In *Semantics and Linguistic Theory*, volume 32, 104–124.
- Kennedy, Christopher. 2015. A “de-Fregean” semantics (and neo-Gricean pragmatics) for modified and unmodified numerals. *Semantics and Pragmatics* 8:10–1.
- Sandoval, Starr. 2023. Ranges: ignorance, temporality, and Ktunaxa. Ms., University of British Columbia.

No vacuous negation in subjunctive questions in Serbian*

NEDA TODOROVIĆ¹
Reed College

1 Introduction

This paper discusses the interaction of negation and subjunctive polar questions in Serbian. To set the stage, consider first indicative polar questions (hereafter IQ). A canonical IQ in Serbian contains *da* (glossed as *da*IND) and a question clitic *li*, as in (1). *Da li* is a stressed form of *li* (Browne 1974, *i.a.*).²

- (1) **Da** **li** deca večeraju?
*da*IND Q kids dine.3PL.PRES
'Are the kids having dinner?' (Oikinomou & Ilić to appear:4)

A subjunctive polar question (hereafter SQ) is shown in (2). Note first that the verbal form in (2) is morphologically the same as the indicative in (1). Second, SQs also contain *da* (glossed here as *da*SUBJV). So, how are SQs different from IQs? They denote modality, despite the absence of an overt modal (Oikinomou & Ilić to appear, henceforth O&I). O&I observe that in (2), the speaker is asking about the addressee's preferences/priorities.

- (2) A: **Da** deca večeraju? B: Ne. 'No'
*da*SUBJV kids dine.3PL.PRES
'Should the kids have dinner?'
(adapted from O&I to appear:5)

* I would like to thank Hotze Rullmann for selflessly sharing his insights with me over the years, which helped me make my arguments stronger and clearer. In addition, his unconditional support had an enormous impact on my growth as a researcher, teacher, and presenter. Dank je wel, Hotze!

¹ For helpful comments and discussion, I would like to thank Lisa Matthewson, Mariia Onoeva, and Mariia Razguliaeva, as well as the audience of the UofT International Workshop on the Semantics of Non-canonical Questions and the attendees of the 2023 EGG summer school. All errors are mine.

² IQs can also be formed, e.g., with *je li*, also a stressed form of *li*, or with a particle *zar*, or through V-fronting. There are subtle pragmatic differences between all those forms. I leave the discussion of these forms aside (see Todorović 2023).

*Da*SUBJV and the present tense form denote modality in non-question contexts as well, e.g., with commands (Kaufmann et al. to appear), as in (3a), or wishes, as in (3b).^{3,4}

- (3) a. **Da** čitaš ovu knjigu!
*da*SUBJV read.2SG.PRES this book
 ‘Read this book (already)!’ (Kaufmann et al. to appear:7)
- b. **Da** ti se sve želje ostvare!
*da*SUBJV you REFL all wishes come.true.3PL.PRES
 ‘May all your wishes come true!’

The IQ in (1) and the SQ in (2) can combine with negation, as shown in (4a) and (4b,c), respectively. In the IQ in (4a), negation is interpretable. In an SQ, the negation can be interpretable, as in (4b), or not, as in (4c). In (4c), the speaker is asking or wondering if the kids are having dinner; the negation seems vacuous. I will label IQs with negation as NegIQ, SQs with contentful negation as NegSQ₁, and SQs with seemingly vacuous negation as NegSQ₂.

- (4) a. **Da** li deca **ne** večeraju?
*da*IND Q kids NEG dine.3PL.PRES
 ‘Are the kids not having dinner?’ (NegIQ)
- b. **Da** deca **ne** večeraju?
*da*SUBJV kids NEG dine.3PL.PRES
 ‘Should the kids not have dinner?’ (NegSQ₁)
- c. **Da** deca (možda) **ne** večeraju?
*da*SUBJV kids maybe NEG dine.3PL.PRES
 ‘Could it be that the kids are having dinner (I wonder)?’
 (NegSQ₂)

³ *Da*SUBJV can occur in certain complements. See Todorović (2015), Todorović and Wurmbrand (2020), and Kaufmann et al. (to appear) for arguments that *da*SUBJV is semantically different from *da* found in indicative complements.

⁴ SQs do not contain *li*. It is, however, possible to combine *da li* and *da*SUBJV, as in (i). O&I report that, unlike (2), (i) has a more introspective flavor; the answer is not required. I leave the discussion of this construction aside.

- (i) A: **Da** li da deca večeraju? (B: Ne. ‘No’)
 A: *da*IND Q *da*SUBJV kids dine.3PL.PRES
 A: ‘Should the kids have dinner (I wonder)?’ (O&I to appear:5)

This is even more evident in (5) with past forms. In the NegIQ in (5a), the negation is interpretable — the speaker is asking whether the kids did not have dinner. In the NegSQ₂ in (5b), the negation seems vacuous — the speaker is asking or wondering whether the kids had dinner. Note that, although the negation can occur in two positions in (5), it is always contentful in the NegIQ in (5a) but not in the NegSQ₂ in (5b).

- (5) a. Da li {nisu} deca {nisu}
*da*IND Q NEG.be.3PL.PRES kids NEG.be.3PL.PRES
 večerala?
 dine.PART.F.SG
 ‘Did the kids not have dinner?’ (NegIQ)
- b. Da {nisu} deca {nisu}
*da*SUBJV NEG.be.3PL.PRES kids NEG.be.3PL.PRES
 večerala?
 dine.PART.F.SG
 ‘Could it be that the kids had dinner (I wonder)?’ (NegSQ₂)

The interpretation of NegSQ₂s immediately raises the question of whether the negation in NegSQs is sometimes vacuous. I argue that it is not. The evidence comes from the interaction between the modal and the negation in NegSQ₂s. I also show that NegSQ₂s are sensitive to epistemic and evidential bias (Sudo 2013), similarly to NegIQs in the language. While SQs have been discussed in the formal literature (O&I to appear), the interaction of SQs and negation has not. This paper aims to contribute to the discussion of the interaction of Slavic polar questions with negation (Abels 2005; Staňková 2023; Zanon 2023) by analyzing novel Serbian data.

The paper is organized as follows. In Section 2, I show that the negation in NegSQ₂s seems vacuous, based on the distribution of Negative Polarity Items (NPIs). However, I argue that the distribution of NPIs is due to the syntax of negation and not due to its vacuous status. In Section 3, I show that the negation in NegSQ₂s cannot be vacuous, based on its interaction with modality. In Section 4, I show that NegSQ₂s are biased, similarly to NegIQs. Section 5 concludes the paper.

2 Negation only seems vacuous

In this section, I discuss the distribution of NPIs in NegSQs. The data

seem to indicate that the negation is vacuous in NegSQ₂. However, I show that the effects arise from syntax, not semantics.

Serbian has two types of NPIs: *ni*-NPIs and *i*-NPIs. *Ni*-NPIs and *i*-NPIs are in complementary distribution (Progovac 1988 et seq.). *Ni*-NPIs need to be licensed by a clausemate negation, as shown for *nikoga* ‘nobody’ in a simple clause in (6a). The embedded *nikoga* in (6b) is licensed by the clausemate negation, but not by the matrix negation.

- (6) a. Jovana ***(ne)** voli **nikoga**.
 Jovana NEG love.3SG.PRES nobody
 ‘Jovana doesn’t like anybody.’
- b. Jovana ***(ne)** kaže [da ***(ne)** voli
 Jovana NEG say.3SG.PRES *da*IND NEG love.3SG.PRES
nikoga].
 nobody
 ‘Jovana says that she doesn’t like anybody.’

I-NPIs are licensed, e.g., in IQs, as shown for *ikoga* ‘anybody’ in (7). They cannot be licensed by a clausemate negation, as in a simple clause in (8a). The embedded *ikoga* in (8b) is not licensed by the embedded negation, but it is licensed by the matrix negation.

- (7) Da li je **ikoga** primetila?
*da*IND Q be.3SG.PRES anybody notice.PART.FEM.SG
 ‘Did she notice anybody?’ (IQ)
- (8) a. *Jovana **ne** voli **ikoga**.
 Jovana NEG love.3SG.PRES nobody
 Intended: ‘Jovana doesn’t like anybody.’
- b. Jovana ***(ne)** kaže [da ***(ne)** voli
 Jovana NEG say.3SG.PRES *da*IND NEG love.3SG.PRES
ikoga].
 anybody
 ‘Jovana doesn’t say that she likes anybody.’

The distribution of NPIs can be used to diagnose the nature of negation in SQs. In the NegSQ₁ in (9), the *ni*-NPI *ništa* ‘nothing’ is licensed, while the *i*-NPI *išta* ‘anything’ is not. The reverse holds in the NegSQ₂s in (10a) and (10b)—the *i*-NPI *išta* ‘anything’ is licensed, while

the *ni*-NPI *ništa* ‘nothing’ is not. In the NegSQ₁ in (9), the negation acts as expected: it licenses the clausemate *ni*-NPI and doesn’t license the *i*-NPI. The apparent problem arises with the NegSQ₂s in (10a) and (10b). In both cases, an *ni*-NPI is infelicitous and an *i*-NPI is felicitous, despite the presence of the clausemate negation. Given that the negation dictates the distribution of NPIs in all the other contexts (6–9), the data in (10) might indicate that the negation in NegSQ₂ is vacuous.

- (9) Da deca **ne** večeraju **ništa** /
*da*SUBJV kids NEG.be.3PL.PRES dine.PART.F.SG nothing /
 * *išta*?
 anything
 ‘Should the kids not eat anything for dinner?’ (NegSQ₁)
- (10) a. Da deca **ne** večeraju * *ništa* /
*da*SUBJV kids NEG.be.3PL.PRES dine.PART.F.SG nothing /
išta?
 anything
 ‘Could it be that the kids are eating something for dinner?’
 (NegSQ₂)
- b. Da {**nisu**} deca {**nisu**}
*da*SUBJV NEG.be.3PL.PRES kids NEG.be.3PL.PRES
 večerala * *ništa* / ***išta***?
 dine.PART.F.SG nothing / anything
 ‘Could it be that the kids were eating something for dinner?’
 (NegSQ₂)

Such a conclusion might, however, be premature. The split we see between NegSQ₁s and NegSQ₂s in terms of NPIs resembles the split we see with NegIQs. First, assume that (i) NPI-licensing happens within the TP in Serbian and that (ii) Serbian has two polarity phrases, one below and one above the TP (Progovac 2005). If the negation is below the TP, it should license clausemate *ni*-NPIs but not *i*-NPIs. This happens in simple clauses, as in (6a), and in the NegSQ₁ in (9). It also happens in IQs with low negation, as in (11a) — only the *ni*-NPI is licit. If the negation is above the TP, it should license *i*-NPIs, but not *ni*-NPIs. This happens with IQs with high negation, as in (11b).

- (11) a. Je l' **nije** *ikog / **nikog**
 JE Q NEG.be.3SG.PRES anybody / nobody
 primetila?
 notice.PART.FEM.SG
 'Did she not notice anybody?' (low NegIQ)
- b. **Nije** li **ikog** / *nikog
 NEG.be.3SG.PRES Q anybody / nobody
 primetila?
 notice.PART.FEM.SG
 'Didn't she notice anybody?' (high NegIQ)
 (adapted from Milićević 2006:5)

Crucially, high negation can still be contentful; the reason why the *ni*-NPI is licensed, but *i*-NPI is not is because the negation is above the TP in (11b) (see Milićević 2006 and Todorović 2023 for additional arguments). In other words, the distribution of the NPIs is due to syntax, not semantics.⁷ Crucially, we cannot exclude this option for the NegSQ_{2S} in (10a) and (10b) — the negation could be contentful, but located above the TP. In that case, the distribution of *ni*-NPIs and *i*-NPIs would similarly be captured by syntax and not by the vacuity of the negation. In other words, the distribution of NPIs is not a strong argument for treating negation as vacuous in NegSQ₂.

3 Negation is not vacuous

In this section, I show that the negation in NegSQs interacts with modality in predictable ways, provided it is analyzed as contentful in both NegSQ_{1S} and NegSQ_{2S}.

Consider first modality. O&I argue that *da*SUBJV in SQs is a modal with prioritizing flavor. First, they show that *da*SUBJV in a wh-SQ as in (12) is obligatory in order to ask about the questionee's prioritizing state; otherwise, the question is an IQ. Second, *da*SUBJV is obligatory for the prioritizing reading in the embedded question in (13); otherwise, the question is an IQ.

⁷ Abels (2005) argues for a similar approach to negation in Russian polar questions (cf. Brown & Franks 1995).

- (12) Šta #(da) deca večeraju?
 what *da*SUBJV kids dine.3PL.PRES
 ‘What should the kids have for dinner?’ (O&I to appear: 5)
- (13) Jovan pita Mariju da li da deca
 Jovan ask.3SG.PRES Marija Q *da*SUBJV children
 večeraju.
 dine.3SG.PRES
 with *da*SUBJV: ‘Jovan is asking Marija whether the kids should
 have dinner.’
 without *da*SUBJV: ‘Jovan is asking Marija whether the kids are
 having dinner.’ (O&I to appear:5)

O&I assume that the modal flavor of *da*SUBJV in embedded contexts depends on the matrix predicate and argue that something similar happens in matrix SQs — the *da*SUBJV modal depends on the speech act Question operator (Q_{op}). *Da*SUBJV acquires a prioritizing flavor in the context of questions, as in (14), where the priorities of the addressee of the speech act event matter (ensured by the e variable). It will also be relevant that O&I treat *da*SUBJV as a weak necessity modal.

- (14) [[SUBJV]] / $Q_{-} =$
 $\lambda f \langle \epsilon, stt \rangle \lambda g \langle \epsilon, stt \rangle \lambda e \lambda q \langle st \rangle \lambda w. \forall w' \in \text{Best}_{PRT}(f, g, e, w) \rightarrow q(w')$

Consider now the SQs in (15). Example (15a) is a non-negated SQ. Example (15b) is a NegSQ₁. Suppose now that (i) the negation is contentful and below the TP in (15b), and (ii) *da*SUBJV is a necessity modal within the TP but above the negation. At the level of the TP, the universal modal scopes over the negation. The reading that we get is ‘It must be the case that the kids are not having dinner’. When we introduce the Q_{op} , the modal gets the prioritizing flavor and the question is asking ‘Should the kids not have dinner?’. This matches the reading in (15b). Consider now the NegSQ₂ in (15c). Assume that *da*SUBJV is the same modal, but the negation is contentful and above the TP. In this syntactic constellation, the negation scopes over the modal. The resulting reading is ‘It is not the case that the kids are necessarily having dinner’. Now, if the negation scopes over the universal quantifier, we standardly expect it to be equivalent to an existential quantifier scoping over the negation. In other words, the reading that we have — ‘It is not the case that the kids are necessarily having dinner’ — can be paraphrased as ‘It is possible

that the kids aren't having dinner'.⁸ At this point, we introduce the Q_{OP}. Importantly, the Q_{OP} scopes over the negation, which scopes over the modal. If locality is required for the modal to acquire the prioritizing flavor in SQs, then this will not be possible in (15c) — the negation occurs between the Q_{OP} and the modal. The Q_{OP} then gets us the reading 'Is it possible that the kids are not having dinner?'.⁹ This is a problem because the paraphrase in (15c) does not correspond to it. But this problem is only apparent.

- (15) a. *Da deca večeraju?*
daSUBJV kids dine.3PL.PRES
 'Should the kids have dinner?'
- b. [Q_{OP} [Modal [Neg...]]]:
Da deca ne večeraju?
daSUBJV kids NEG dine.3PL.PRES
 'Should the kids not have dinner?' (NegSQ₁)
- c. [Q_{OP} [Neg [Mod...]]]:
Da deca (možda) ne večeraju?
daSUBJV kids maybe NEG dine.3PL.PRES
 'Could it be that the kids are having dinner (I wonder)?'
 (NegSQ₂)

Let's think about questions — they introduce a set of possible answers (Hamblin 1973, *i.a.*). For the question in (15c) the set of answers would contain the two propositions 'It is possible that the kids aren't having dinner' and 'It is not possible that the kids aren't having dinner' (which can further be paraphrased as 'It must be the case that the kids are having dinner'). Semantically, both options are valid answers for (15c). But pragmatically, the speaker in (15c) might have a slight expectation for the positive answer, i.e., that the kids are having dinner. Thus, while semantically it is possible to ask about either the negative or the positive option, I propose that the positive speaker's bias is what affects the interpretation of this question. So, the semantics of NegSQ₂ is unchanged, showing the scopal interactions of modal, negation, and Q_{OP}.

⁸ I would like to thank Mariia Razguliaeva for sharing her ideas on this matter.

⁹ NegSQ_{2S} with past forms show the same pattern as NegSQ_{2S} with present forms.

In the following section, I show that NegSQ₂s are indeed sensitive to speaker's bias. In that respect, they are parallel to negative IQs in Serbian.

4 NegSQ₂s are contextually-sensitive, just like NegIQs

In this section, I show that NegSQ₂s are sensitive to contextual information, which captures the speaker's expectations of what the answer should be. In terms of being contextually sensitive, NegSQ₂s align with NegIQs in Serbian.

The interaction of negation and questions in different contexts has been extensively studied across languages (Büring & Gunlogson 2000, *i.a.*). Negative questions tend to be biased (Ladd 1981), and Sudo (2013), for example, identifies that the bias is epistemic (stemming from the speaker's beliefs) or evidential (stemming from the context) (this is by no means an exhaustive list). Sudo also argues that different values of epistemic or evidential bias capture the differences in the distribution of questions with high and low negation.

To illustrate this with Serbian IQs, when there is no bias, a positive IQ as in (16a) is felicitous, but a low NegIQ as in (16b) or a high NegIQ as in (16c) is not.

(16) *Context: Your roommate was at her friend Milana's birthday party. One typically makes a birthday cake for that occasion, but you don't know if Milana does that too. You also don't know if she made a cake this time. You ask your roommate:*

- a. Da li je Milana pravila tortu?
*da*IND Q be.3SG.PRES Milana make.PART.F.SG cake
 'Did Milana make a cake?' (IQ)
- b. #Je l' Milana **nije** pravila tortu?
JE Q Milana NEG.be.3SG.PRES make.PART.F.SG cake
 'Did Milana not make a cake?' (low NegIQ)
- c. #**Nije** li Milana pravila tortu?
 NEG.be.3SG.PRES Q Milana make.PART.F.SG cake
 'Didn't Milana make a cake?' (high NegIQ)

Manipulating the value of the bias further affects the distribution of negative questions: while both low and high NegIQs are felicitous with

positive epistemic and negative evidential bias in (17), only low NegIQ is felicitous with neutral epistemic and negative evidential bias in (18) (see Todorović 2023 for a complete distribution of these questions in context).

(17) *Context: Your roommate was at Milana's birthday party. One typically makes a birthday cake for that occasion. You know that Milana likes making cakes and you think she made one this time as well. Whenever there's cake at a party, your roommate brings you the leftovers. You open the fridge, but don't see any cake leftovers. You ask your roommate:*

- a. # Da li je Milana pravila tortu?
*da*IND Q be.3SG.PRES Milana make.PART.F.SG cake
 'Did Milana make a cake?' (IQ)
- b. Je l' Milana **nije** pravila tortu?
 JE Q Milana NEG.be.3SG.PRES make.PART.F.SG cake
 'Did Milana not make a cake?' (low NegIQ)
- c. **Nije** li Milana pravila tortu?
 NEG.be.3SG.PRES Q Milana make.PART.F.SG cake
 'Didn't Milana make a cake?' (high NegIQ)

(18) *Context: Your roommate was at Milana's birthday party. One typically makes a birthday cake for that occasion, but you don't know if Milana does that too. Whenever there's a cake at some party, your roommate brings you the leftovers. You open the fridge, but don't see any cake leftovers. You ask your roommate:*

- a. # Da li je Milana pravila tortu?
*da*IND Q be.3SG.PRES Milana make.PART.F.SG cake
 'Did Milana make a cake?' (IQ)
- b. Je l' Milana **nije** pravila tortu?
 JE Q Milana NEG.be.3SG.PRES make.PART.F.SG cake
 'Did Milana not make a cake?' (low NegIQ)
- c. # **Nije** li Milana pravila tortu?
 NEG.be.3SG.PRES Q Milana make.PART.F.SG cake
 'Didn't Milana make a cake?' (high NegIQ)

NegSQ₂s, like NegIQs, are also sensitive to epistemic and evidential bias. First, they cannot be used in a neutral context, as shown in (19).

- (19) *Context: Your roommate was at her friend Milana's birthday party. One typically makes a birthday cake for that occasion, but you don't know if Milana does that too. You also don't know if she made a cake this time. You ask your roommate:*

# Da	nije	Milana pravila	tortu?
	<i>da</i> SUBJV	NEG.be.3SG.PRES Milana make.PART.F.SG	cake
		'Could it be that Milana made a cake?'	(NegSQ ₂)

Second, NegSQ₂s require neutral or positive evidential bias; the latter is shown in (20).

- (20) *Context: Your roommate was at her friend Milana's birthday party. One typically makes a birthday cake for that occasion. You don't know if Milana made a cake. But you know that Milana likes making cakes and you think that she made one this time as well. You ask your roommate:*

Da	nije	Milana pravila	tortu?
	<i>da</i> SUBJV	NEG.be.3SG.PRES Milana make.PART.F.SG	cake
		'Could it be that Milana made a cake?'	(NegSQ ₂)

The full distribution is shown in Table 1. The distribution of NegSQ₂s does not exactly match the distribution of NegIQs. Yet, even low and high NegIQs do not match in their distribution, as shown, e.g., in (18). What both NegIQs and NegSQ₂s, however, have in common is that they show either epistemic or evidential bias. This aligns with negative questions across languages.

Table 1: Distribution of NegSQ₂s in Serbian

		Epistemic	
Evidential		positive	neutral
	positive		NegSQ ₂
	neutral	NegSQ ₂	NegSQ ₂ (ironic)
	negative		

One way to capture the distribution of these questions would be along the lines of AnderBois' (2019) inquisitive semantics approach to American English negative polar questions. This approach assumes that certain expressions, e.g., existential quantifiers and disjunction, introduce alternatives, i.e., the inquisitive part of an expression that can serve as a starting point for further conversation. Negation, on the other hand, gets rid of any alternatives. In AnderBois' approach, two syntactic positions of negation (above or below the TP) and the distribution of the content of negation (universal quantifier and complementation) between the two heads, affects what will be highlighted as the prominent issue for further discussion. In other words, different syntax and semantics of negation in low and high negation questions will affect which issues are relevant for further discussion; this would reflect different biases of the speaker. Todorović (2023) shows that the same can be applied to Serbian NegIQs. As for the NegSQ₂s, one might expect them to show similarities with high NegIQs, given their syntax. However, they do not match — high NegIQs are restricted to contexts with positive epistemic and negative evidential bias. One of the reasons for their differences might be that the syntax-semantics of the modal in NegSQ₂s also plays a role in highlighting certain issues and reflecting a particular bias. I leave this issue for further research. Importantly, NegSQ₂s, like NegIQs, cannot be used in a neutral context, but are sensitive to contextual information.

5 Conclusion

In certain NegSQs in Serbian, NegSQ₁s, the negation is contentful, while in the others, NegSQ₂s, it seems to be vacuous. Such a division seems to find support in syntax, since NegSQ₁ licenses NPIs and NegSQ₂ does not. However, I have argued that the differences that we see are not due to their different semantics. With NegSQ₁s, the negation is in a local enough relationship with the NPI to license it, while with NegSQ₂s, it is too far from the NPI to license it. In other words, the semantics of negation is the same, but the syntax affects the NPI licensing options. I further argued that the differences in syntax of negation in NegSQ₁s and NegSQ₂s also affect the scopal interactions between the modal and negation, deriving different interpretations. In either case, the patterns can be captured just in case the negation is contentful. Finally, I showed that NegSQ₂s show speaker's bias in terms of which answer they expect to hear. In that respect, NegSQ₂s are not different from NegIQs. While

there is more to be said about the properties of NegSQs, it seems the negation in them is not different from negation elsewhere in the language.

References

- Abels, Klaus. 2005. "Expletive Negation" in Russian: A Conspiracy Theory. *Journal of Slavic Linguistics* 13(1):5–7.
- AnderBois, Scott. 2019. Negation, alternatives, and negative polar questions in American English. In Malte Zimmermann, Klaus von Heusinger, & Edgar Onea (eds.), *Questions in Discourse*. Leiden, Netherlands: Brill, 118–171.
- Brown, Sue & Steven Franks. 1995. Asymmetries in the Scope of Russian Negation. *Journal of Slavic Linguistics* 3(2):239–287.
- Browne, Wayles. 1974. On the problem of clitic placement in Serbo-Croatian. In R. D. Brecht and Catherine Chvany (eds.), *Slavic transformational syntax*. Ann Arbor, MI: University of Michigan, 36–52.
- Büring, Daniel & Christine Gunlogson. 2000. *Aren't positive and negative polar questions the same?* Ms. University of California, Los Angeles & University of California, Santa Cruz.
- Hamblin, Charles L. 1973. Questions in Montague English. *Foundations of Language* 10:41–53.
- Kaufmann, Magdalena, Neda Todorović & Ivana Jovović. To appear. Obviate me (not): Obviation effects in Serbian main and embedded clauses. Special issue of *Journal of Slavic Linguistics: Proceedings of FASL 30*.
- Ladd, D. Robert. 1981. A first look at the semantics and pragmatics of negative questions and tag questions. *Chicago Linguistic Society* 17: 164–171.
- Milićević, Nataša. 2006. On negation in yes/no questions in Serbian. In *UiL OTS Working Papers*, 29–47.
- Oikonomou, Despina & Ivona Ilić. To appear. Subjunctive Questions in Serbian. Special issue of *Journal of Slavic Linguistics: Proceedings of FASL 30*.

- Progovac, Ljiljana. 1988. *A Binding Approach to Polarity Sensitivity*, Ph.D. dissertation, University of Southern California, Los Angeles, CA.
- Progovac, Ljiljana. 2005. *A Syntax of Serbian*. Bloomington, IN: Slavica Publishers.
- Staňková, Anna. 2023. *The expression of speaker's bias in Czech polar questions*. M.A. thesis, University of Prague.
- Sudo, Yasutada. 2013. Biased polar questions in English and Japanese. In Daniel Gutzmann, & Hans-Martin Gärtner (eds.), *Beyond Expressives: Explorations in Use-Conditional Meaning*. Leiden, Netherlands: Brill, 275–295.
- Todorović, Neda. 2015. Tense and aspect (in)compatibility in Serbian matrix and subordinate clauses. *Lingua* 167:82–111.
- Todorović, Neda & Susi Wurmbrand. 2020. Finiteness across domains. In Teodora Radeva-Bork & Peter Kosta (eds.), *Current Developments in Slavic Linguistics: Twenty Years After*. Frankfurt am Main, Germany: Peter Lang, 47–66.
- Todorović, Neda. 2023. Negation and questions in Serbian. Lecture slides, *EGG Summer School*, July 24–31, Novi Sad.
- Zanon, Ksenia. 2023. “Expletive Negation” revisited: On some properties of negated polar interrogatives in Russian. Paper presented at *FASL 32*. University of Indiana, Bloomington.

A maximality – exhaustivity connection: The semantics and pragmatics of *-dake* in Japanese*

SATOSHI TOMIOKA
University of Delaware

The topic of this paper is the meaning of the Japanese expression *dake*, which is most commonly translated as *only* in English. The meaning of the English *only* is by no means simple or uncontroversial.¹ My story of *dake* is also rather complicated, but it is not a direct consequence of inheriting the known complexity of *only*. On the contrary, I argue that the semantic core of *dake* is fundamentally different from that of *only*. Despite the difference, however, the two expressions are often functionally equivalent and can describe the same state of affairs. The common practice of equating *dake* to *only* is undoubtedly based on this practical aspect, but we should be reminded that functional equivalence does not necessarily mean semantic equivalence, as there are some known cases

* Hotze was three years ahead of me at UMass, which means that we only had two years together in the graduate program, but I learned a lot from him during those two years. What I appreciated (and still appreciate) the most about Hotze is his unbiased, non-dogmatic attitude towards other people's ideas. I always felt secure enough to try out my (frequently silly) ideas on him, and it is quite admirable that he always managed to find something useful to say. I didn't fully grasp the impact of Hotze's dissertation until a few years after his graduation, and this paper is a very delayed demonstration of my appreciation of his work. This research was in part supported by The JSPS Core-to-Core Program, A. Advanced Research Networks "International Research Network for the Human Language Faculty" (#JPJSCCAJ221702004)

¹ In Horn (1969), it is argued that *only X is Y* asserts that no non-X is Y while the 'prejacent' proposition that X is Y is presupposed. While Horn's highly influential analysis has nonetheless been challenged, the debate seems to focus on the presuppositional status of the prejacent (e.g., van Rooji and Schulz 2005, Ippolito 2008). On the other hand, the asserted or 'at-issue' nature of the exhaustive meaning is rarely questioned. One notable exception is Zeevat (2009), whose proposal will be discussed in detail later.

of such discrepancies. For instance, a definite plural and a universally quantified plural (e.g., *the students* vs. *all the students*) arguably have different semantic denotations that are nonetheless hard to distinguish in terms of truth conditions (cf. Brisson 1998). In Dayal's (2000) analysis of the Wh-scope marking structure in Hindi, the scope marking strategy is not semantically identical to the overtly wh-moved counterpart, but the two strategies are not easily distinguishable in terms of their communicative function. The current study argues that the *dake-only* comparison presents another case of this kind.

Let us begin with the informal observation that an English sentence with *only* can have two different translations in Japanese.

- (1) Only Mary passed the exam.
- (2) a. Mari-dake-ga siken-ni ukatta.
Mari-DAKE-Nom exam-DAT PASS.PAST
b. Mari-sika siken-ni ukar-anakatta
Mari-SIKA exam-DAT PASS-NEG.PAST

The truth conditions of the two sentences in (2) have the two components: (i) Mari passed the exam, and (ii) no other relevant people did. (2b) involves the negative concord item *XP-sika*, which mimics patterns found in other languages, such as *ne ... que XP* in French. It is sometimes regarded as a type of exceptive construction, and its interpretation, 'nothing/no one except for XP', is practically identical to that of *only*. I do not have much more to say about the meaning of *sika...nai*, however. I will assume that it has the same semantic content as *only*. Although the two versions of *only* in Japanese can describe the same situation, their interchangeability breaks down in some contexts, as Kuno (1999) and Yoshimura (2007) discovered. According to these authors, their differences boil down to the strength of their negative meaning: Informally speaking, *dake* generates weaker negative meaning than *-sika...nai* does. For instance, consider the following English example.

- (3) Q: Why didn't Daisuke get the job?
A: Because he only speaks Japanese.

The felicity of this mini-discourse shows that the negative proposition, 'Daisuke does not speak any languages other than Japanese', is readily

available with *only*. When we compare the two Japanese expressions of ‘only’, however, an unexpected pattern emerges.

- (4) Q: Why didn’t Daisuke get that job?
- a. ??? nihongo-dake hanas-eru-kara-desu.
Japanese-DAKE speak-can-because-be
Intended: ‘Because he can speak only Japanese.’
 - b. nihongo-sika hanas-e-nai-kara-desu.
Japanese-SIKA speak-can-Neg-because-be
‘Because he cannot speak any languages other than Japanese.’
 - c. hanas-eru-no-ga nihongo-dake-da-kara-desu.
speak-can-NML-Nom Japanese-DAKE-be-because-be
‘Because Japanese is the only language that he can speak.’

Surprisingly, the use of *dake* is rather inadequate in the context above although it is perhaps not outright infelicitous. In this discourse context, *sika...nai* is a better choice, as it can more effectively communicate Daisuke’s inability to speak other languages. Interestingly, however, *dake* becomes much more acceptable when it is clefted, as is demonstrated in (4c).

Another environment in which *dake* and *sikanai* behave differently is a conditional sentence. Kuno (1999) notes that an *if*-clause that is interpreted as ‘as long as’ can embed *dake*, but not *sika...nai*.

- (5) sekai-ryokou-o suru-niwa,
world-travel-ACC do-in.order.to
‘In order to make an around-the-world trip’
- a. eego-dake hanas-er-eba ii.
English-DAKE speak-can-if good
‘it is all right as long as (you) can speak English.’
 - b. #eego-sika hanas-e-nak-ereba ii.
English-SIKA speak-can-NEG-if good
‘it is all right as long as (you) cannot speak any other languages besides English.’

- c. #hanas-eru-no-ga eego-dake deare-ba ii.
 speak-can-NML-NOM English-DAKE be-if good
 ‘it is all right as long as it is only English that you can
 speak.’

The exhaustivity meaning of *dake* in (5a) is weakened to the extent that its contribution is almost invisible. The *-sika...nai* counterpart is pragmatically odd, which is expected since its English translation with *only* is equally infelicitous in the same context. Once again, the clefted *dake* patterns with *-sika...nai*. Previously, Kuno (1999) suggested that the positive meaning (= the prejacent) is the primary meaning of the sentence with *dake*, and the negative meaning (= the exhaustive meaning) the secondary. Yoshimura (2007), on the other hand, argued that the positive meaning is ‘asserted’, while the negative meaning is ‘entailed’ in the sense of Horn (2002). However, neither author considers the cleft data, and it is unclear at best how the clefted *dake* can elicit the negative meaning comparable to *sika...nai*.²

In this paper, I offer an alternative analysis in which the exhaustivity meaning (= the negative quantification over non-weaker alternatives) is altogether absent in *dake*. I argue that the exhaustive-like meaning of *dake* is rooted in its use as a degree expression, roughly paraphrased as ‘as much/many as’, ‘the upper limit’. In particular, it inherits the notion of ‘maximality’, which is common in degree expressions (cf. von Stechow 1984, Rullmann 1995) and the exhaustive interpretation is inferable from it and the additional ‘mirative’ import (cf. Zeevat 2009). Let us begin with Futagi’s (2004) observation that *dake*, which was historically derived from *take* ‘length, height, limit’, can still be used as a degree expression in the contemporary Japanese, as shown below.³

² There are a couple of relevant papers that I unfortunately cannot include in the discussion here: Ido and Kubota (2021) and Oshima (2023).

³ In addition, there are other degree/scale expressions that come close to the meaning of ‘only’: *bakari*, which can alternatively mean ‘approximately’. This morpheme derives from the verb *hakaru* ‘to measure’. *X-kagiri* can also mean ‘only X’, and *kagiri* is a nominal form of the verb *kagiru* ‘to limit’. Below are examples of *bakari*.

- (i) a. soko-no niku-o 1-kiro-bakari kudasai.
 there-GEN meat-ACC 1-kilo-BAKARI give.me.please

- (6) a. Ringo-o aru-dake/ari-ttake motte-ki-ta.
apple-ACC exist-DAKE/exist-DAKE have-come-Past
'(I) brought as many apples as I had.'
- b. 5-en kitte-o 100-en-dake kudasai.
5-yen stamp-ACC 100-yen-DAKE give.me.please
'Please give [me] one hundred yen worth of five-yen stamps.'
= Futagi (2004, (222a))
- c. Hikkosi-ni dore-dake okane-ga kakarimasita-ka?
moving-DAT which-DAKE money-NOM cost.PAST-Q
'How much money did it cost (you) to move?'

My proposal is based on the intuition that *dake* maintains the connection to the degree meaning even when it is used as an exhaustive/exclusive particle. I argue that *X-dake* means something similar to 'as much/many as X,' 'X is the upper limit', or 'to the extent of X'. When we make reference to degrees, the maximality operation is often needed. For instance, comparatives require maximality (cf. von Stechow 1984, Rullmann 1995, Schwarzschild 2008).

- (7) a. Anna is taller than Maria is.
b. There is a degree *d* such that Anna is *d*-tall and *d* is higher than the maximal degree of Maria's height.
c. Definition of the Maximality Operator *max*, Rullmann (1995, (21)):
Let DEG be a set of degrees ordered by the relation \leq , then
 $max(DEG) = \iota d [d \in DEG \wedge \forall d' \in DEG [d' \leq d]]$.

The effect of maximality with *dake* is straightforwardly interpreted in (6c):

'Please give me one kilo of that meat over there (but you needn't be exact).'

- b. aitsu-wa niku-bakari-tabete, yasai-wa zenzen tabe-nai.
that.guy-TOP meat-BAKARI-eat, vegetable-TOP at.all eat-NEG
'(Every time I see him eat), that guy only eats meat and never touches vegetables.'

- (8) a. LF of (6c): [$_{CP}$ dore-dake [λd_1 [$_{IP}$ hikkosi-ni t_1 okane-ga kakarimashita]] -ka]
 b. The meaning: $\{p: \exists d \wedge p = [d = \max(\lambda d'. \text{it cost (you) } d' \text{-much money to move})]\}$

In this example, the complement of *dore-dake* ‘how much’ denotes a set of degrees, and *max* chooses the maximal degree out of that set. In applying the maximization process to the exhaustive/exclusive use of *dake*, I maintain that the maximization applies to the complement of *XP-dake* but also make the following transitional steps: (i) To avoid complications, I focus on cases of *XP-dake*, where *XP* is an entity denoting expression, (ii) the complement of *XP-dake* denotes a set of entities, and (iii) the maximization applies to the complement, yielding the maximal entity. The last step is practically identical to the semantics of definite plurals. In most cases in which degree expressions are involved, the maximization operation takes place implicitly, but I hypothesize that *dake* lexically encodes the maximization. Under this scheme, the denotation of *dake* is (9a). Unlike (7b), *max* in (9a) does not take a set of degrees but a set of entities instead.

- (9) a. $\llbracket \text{dake} \rrbracket = \lambda x. \lambda P. \max(P) = x$
 b. Let *P* be a set of atomic and plural entities, then
 $\max(P) = \iota x [x \in P \wedge \forall y \in P [y \leq x]]$

With this meaning of *dake*, the sentence (10a) is interpreted as (10b).

- (10) a. Aya-to Saki-to Eri-dake-ga ukat-ta.
 Aya-and Saki-and Eri-DAKE-Nom PASS-PAST
 b. The maximal individuals who passed are Aya, Saki and Eri.
 \approx The people who passed are Aya, Saki, and Eri.

While (10b) describes the meaning of (10a) fairly well, something is amiss. The same discomfort is felt when we compare the two near-equivalent English sentences, *only A, B, and C passed* and *the people who passed are A, B, and C*. The definite plural paraphrase (10b) is compatible with a situation where Aya, Saki and Eri are all the students who took the exam. Clearly, however, (10a) is utterly inappropriate in such a situation, just as is the case with the English sentence with *only*. One remedy is to appeal to the Roothian focus semantic import here (Rooth 1992): Let the

NP, *Aya*, *Saki*, and *Eri* be focused and generate a non-singleton set as its focus value. Then, the existence of some entities other than three individuals must be included.

Alternatively, we can assume, following Zeevat's (2009) analysis of *only* in English, that there is a mirative (surprise) component in sentences with *dake*: The argument of *dake* was less than expected or lower than some salient standard. The following is a brief summary of Zeevat's analysis of mirativity of focus-sensitive particles.

- (11) a. Some particles add 'mirative/surprise' flavors. *even*: more than expected, *only*: less than expected, *already*: earlier than expected, *still*: later than expected, etc.
- b. The exhaustivity meaning in a sentence with *only* is derivable with focus; Even without *only*, the sentence has the exhaustive meaning, as it is typically considered as the complete (exhaustive) answer to a QUD.
- c. Thus, the mirativity is the sole meaning of *only*, and it is regarded as not-at-issue (a weak presupposition in Zeevat's term).

With the added mirativity, the meaning of (10a) is (12).

- (12) a. The maximal individuals who passed are *Aya*, *Saki* and *Eri*.
- b. The maximal individual (*Aya* & *Saki* & *Eri*) was less/fewer than expected.

(12b) cannot be satisfied unless there are others who took the test, could have also passed but didn't. As a result, the maximality becomes very close to, and practically indistinguishable from, the exhaustive meaning of *only*. Since the mirative meaning is supposed to be not-at-issue, however, the exhaustivity generated by *dake* is expected to be weaker than the English *only*.

This way of thinking the (non-)exhaustivity of *dake* paves the way to explain why the clefted *dake* becomes more like *sika...nai*. First of all, it should be noted that the cleft construction itself can generate exhaustivity. In the context where several students took the exam, uttering *It is Eri that passed the exam* means that only *Eri* passed the exam. However, the exhaustivity of a cleft sentence does not always match that of *only*. Consider the Japanese examples below.

(13) Did both Aya and Eri pass the exam?

lie, ukat-ta no-wa Eri-#(dake)-desu.

No, pass-PAST NML-TOP Eri-(DAKE)-be

‘It is #(only) Eri that passed the exam.’

The cleft construction does not have suitable exhaustivity for a negative answer to a ‘both’ question, and in such a situation, the addition of *dake* is necessary. To the extent that the English translation shows the same pattern, (13) may not be surprising. However, I have hypothesized that the meaning of *dake* is not the same as the English *only*. Then, why does the addition of *dake* elicit the same effect as *only*? Putting the puzzle slightly differently, we have witnessed that the clefted *dake* seems to elicit the kind of exhaustivity or negativity that is comparable to *sika...nai*. How does this strengthened exhaustivity come about? First, the cleft construction involves focus, and that is undeniable. When a clefted *X-dake* phrase is interpreted as exhaustive as *only* or *sika...nai*, what is actually focused is the particle *dake*, rather than *X* or the whole *X-dake*, which is indicated by the prosody. In the sentence (13), the focal accent is placed on the particle *dake* alone. I hypothesize that this focus pattern leads to the generation of a polar alternative, as shown below.

(14) a. ukat-ta no-wa Eri-DAKE-desu.

pass-PAST NML-TOP Eri-DAKE-be

b. {Eri is the maximal individual who passed, Eri is not the maximal individual who passed}

Before discussing how these alternatives are made use of, it is necessary to examine the second alternative: Eri is not the maximal individual who passed. Technically speaking, this sentence is true either if Eri and someone else passed or if Eri herself did not pass in the first place. However, there are good indications that the sentence (14a) presupposes that Eri passed. For instance, the negation and the polar question formation tests show that Eri’s passing is presupposed.

(15) a. Ukat-ta no-wa Eri-DAKE-dewa ari-masen.

pass-PAST NML-TOP Eri-DAKE-be exist-NEG

‘It is not only Eri that passed.’ \rightsquigarrow Erika passed.

- b. Ukat-ta no-wa Eri-DAKE-desu-ka?
 pass-PAST NML-TOP Eri-DAKE-be-Q
 ‘Is it only Eri that passed? \rightsquigarrow Erika passed.

The second alternative, therefore, means that Eri and someone else passed. The exhaustivity associated with the cleft construction negates this alternative. Suppose that there are three exam takers, Aya, Saki, and Eri. The proposition that Eri is not the maximal individual who passed is equivalent to (16):

- (16) Aya and Eri passed \vee Saki and Eri passed \vee Aya, Saki and Eri passed

Combined with the presupposition that Eri passed, the negation of (16) leads to the negation of the two alternatives: Aya didn’t pass, and Saki didn’t pass. This is precisely the same semantic effect of *only/sika...nai*, which involves the negation of all the non-weaker alternatives.

The unexpected behavior of *dake* in conditionals is also accounted for. Recall:

- (17) sekai-ryokou-o suru-niwa,
 world-travel-ACC do-in.order.to
 ‘In order to make an around-the-world trip’
- a. eego-dake hanas-er-eba ii.
 English-DAKE speak-can-if good
 ‘it is all right as long as (you) can speak English.’
- b. # eego-sika hanas-e-nak-ereba ii.
 English-SIKA speak-can-NEG-if good
 ‘it is all right as long as (you) cannot speak any other languages besides English.’
- c. # hanas-eru-no-ga eego-dake deare-ba ii.
 speak-can-NML-NOM English-DAKE be-if good
 ‘it is all right as long as it is only English that you can speak.’

The *as-long-as* interpretation is often elicited when the consequence has such expressions as *good (enough, sufficient, X is content, etc.)*. In such a conditional, the antecedent *p* provides some ‘minimally sufficient’ criterion for the consequent clause to hold. The notion of ‘minimally suffi-

cient' can be defined in semantic terms (as is predicted by propositional logic), but it can be based on something more pragmatic.

- (18) If you can bring fruits salad, that will be sufficient.
- a. You need not bring anything in addition to fruits salad.
 - b. You need not bring anything fancier / more complicated to make.

Turning to the felicitous *dake* sentence (17a), we can easily imagine the following ordering based on the number of languages that the addressee can speak.

- (19) $\max(\lambda x. \text{is-able-to-speak}(x)(\text{addressee})) = \text{English} <$
 $\max(\lambda x. \text{is-able-to-speak}(x)(\text{addressee})) = \text{English} \oplus \text{Spanish}, <$
 $\max(\lambda x. \text{is-able-to-speak}(x)(\text{addressee})) = \text{English} \oplus \text{Spanish} \oplus$
 $\text{Russian} < \dots$

The propositional content of the *if*-clause sets the minimal criterion in this ordering. Therefore, the conditional statement is felicitous.

In contrast, *sika...nai* is infelicitous in this context, and I suggest that the infelicity is predicted because the ordering is reversed. *Sika...nai* primarily asserts the negation of the alternatives. Thus, the ordering is based on the number of languages that the addressee cannot speak. The following is the ordering of possible propositions contained in *if* that are compatible with the prejacent (you speak English). Imagine that there are three relevant languages; English, Russian and Spanish.

- (20) $\text{is-able-to-speak}(\text{Eng} \oplus \text{Rus} \oplus \text{Sp})(\text{addressee}) \ \&$
 $\neg \exists x. \neg \text{is-able-to-speak}(x)(\text{addressee}) <$
 $\text{is-able-to-speak}(\text{Eng} \oplus \text{Sp})(\text{addressee}) \ \&$
 $\neg \text{is-able-to-speak}(\text{Rus})(\text{addressee}),$
 $\text{is-able-to-speak}(\text{Eng} \oplus \text{Rus})(\text{addressee}) \ \&$
 $\neg \text{is-able-to-speak}(\text{Sp})(\text{addressee}) <$
 $\text{is-able-to-speak}(\text{Eng})(\text{addressee}) \ \&$
 $\neg \text{is-able-to-speak}(\text{Rus} \oplus \text{Sp})(\text{addressee})$

The trouble here is that the one that is actually said *you don't speak any languages except for English* is the highest in the ordering. In other words,

it creates a pragmatic anomaly since the highest in the ordering is presented as the minimally sufficient condition. Since the clefted *dake* conveys the exhaustive meaning comparable to *sika..nai*, the ordering is the same as *sika..nai*. Thus, the use of the clefted *dake* is infelicitous as well.

One interesting consequence of the proposed semantics of *dake* is that *dake* is not a focus-sensitive expression if the mirative component is responsible for eliciting (weak) exhaustivity. The idea advocated by Rooth (1985) treats *only* as a focus-sensitive operator. It must have a focus within its scope since its lexical meaning requires a non-singleton set of alternatives. That is not the case with *dake*. The meaning of *dake* spelled out in (12) makes no reference to focus semantic values. Like any other expression, a *dake*-phrase can be focused but need not be. This aspect of *dake* seems to get support from what has come to be known as ‘LF intervention effects’. It has been observed that wh-phrases cannot outscope a certain class of interveners that c-command the wh-phrases at the surface level (cf. Hoji 1985, Kim 2002, Beck 2006, Beck and Kim 2006, Tomioka 2007b, among others). Kim (2002) and Beck (2006) identify potential interveners as focus-sensitive expressions. Tomioka (2007b) uses a more pragmatic notion of focus, but focus-sensitive expressions such as *only*-phrases are considered good candidates for interveners because they tend to be pragmatic foci as well. Surprisingly, however, the experimental studies reported in Kitagawa, Tamaoka and Tomioka (2013) found that *dake*-phrases do not induce intervention effects at all. The following sentences form a minimal pair.

- (21) a. Mariko-dake-ga dare-o sasot-ta-no?
 Mariko-DAKE-NOM who-ACC invite-PAST-Q
 ‘Who did only Mariko invite?’
 b. Dare-o Mariko-dake-ga sasot-ta-no?
 who-ACC Mario-DAKE-NOM invite-PAST-Q
 ‘Who did only Mariko invite?’

In Tomioka (2007a), *dake* is labeled as a weak intervener, and a sentence similar to (21a) is judged questionable while its scrambled counterpart is acceptable. However, this judgment was not reproduced in Kitagawa, Tamaoka and Tomioka’s results: The changing of the c-command relation between a *dake* and a wh-phrase has no effects. The pair of sentences in (21) received acceptability scores that are not significantly dif-

ferent from each other (2.76 vs. 2.88 in the 6 point scale, $p = 0.651$). This result is no longer surprising if the current proposal is right. *Dake* is not exhaustive, and it does not require a focus semantic value for its meaning to be computed.

To sum up, I have argued, going against the conventional wisdom, that

- (22) a. *Dake* encodes maximality, which is rooted in its other self as a degree expression.
- b. The maximality of *dake*, combined with the mirative import, leads to the exhaustive(-like) interpretation, but it is expectedly weak, as it has no explicit negation of non-weaker alternatives.
- c. When clefted, *dake* is contrasted with its ‘non-maximal’ alternative, and the denial of the latter due to the exhaustivity of the cleft construction makes the clefted *dake* more like the true exhaustive *only*.
- d. All these effects can be obtained without making any reference to focus semantic values. Therefore, *dake* is not a focus-sensitive operator.

References

- Beck, Sigrid. 2006. Intervention effects follow from focus interpretation. *Natural Language Semantics* 14:1–56.
- Beck, Sigrid, and Shin-Sook Kim. 2006. Intervention effects in alternative questions. *Journal of Comparative Germanic Linguistics* 9:165–208.
- Brisson, Christine. 1998. Distributivity, maximality and floating quantifiers. Doctoral Dissertation, Rutgers University.
- Dayal, Veneeta. 2000. Scope marking: Cross-linguistic variation in indirect dependency. In *Wh-scope marking*, ed. Uli Lutz, Gereon Müller, and Arnim von Stechow, 157–193. Amsterdam: John Benjamins.
- Futagi, Yoko. 2004. Japanese focus particles at the syntax–semantics interface. Doctoral Dissertation, Rutgers University.

- Hoji, Hajime. 1985. Logical form constraints and configurational structures in Japanese. Doctoral Dissertation, University of Washington.
- Horn, Laurence. 1969. A presuppositional analysis of *only* and *even*. In *Papers from the fifth Regional Meeting of the Chicago Linguistic Society*, 97–108.
- Horn, Laurence R. 2002. Assertoric inertia and NPI licensing. In *Papers from the 38th annual meeting of the Chicago Linguistic Society, parasession on polarity and negation*, 55–82. CLS.
- Ido, Misato, and Yusuke Kubota. 2021. The hidden side of exclusive focus particles: An analysis of *dake* and *sika* in Japanese. *GENGO KENKYU (Journal of the Linguistic Society of Japan)* 160:183–213.
- Ippolito, Michela. 2008. On the meaning of *only*. *Journal of Semantics* 25:45–91.
- Kim, Shin-Sook. 2002. Intervention effects are focus effects. *Japanese/Korean Linguistics* 10:615–628.
- Kitagawa, Yoshihisa, Katsuo Tamaoka, and Satoshi Tomioka. 2013. Prosodic matters in intervention effects in Japanese: An experimental study. *Lingua* 124:41–63.
- Kuno, Susumu. 1999. The syntax and semantics of the *dake* and *sika* constructions. In *Harvard working papers in linguistics*, ed. Bert Vaux and Susumu Kuno, volume 7, 144–172. Harvard University.
- Oshima, David Y. 2023. Semantic variation in exclusive quantifiers: English *only*, Japanese *dake*, *dake-wa*, and *shika*, and the cleft construction. *Natural Language & Linguistic Theory* 1–33.
- van Rooij, Robert, and Katrin Schulz. 2005. Only: Meaning and implicatures – the very incomplete short version. In *Proceedings of SuB 9*, ed. Emar Maier, Corien Bary, and Janneke Huitink, 314–324.
- Rooth, Mats. 1985. Association with focus. Doctoral Dissertation, University of Massachusetts, Amherst.
- Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1:75–116.
- Rullmann, Hotze. 1995. Maximality in the semantics of *wh*-constructions. Doctoral Dissertation, University of Massachusetts at Amherst.

- Schwarzschild, Roger. 2008. The semantics of comparatives and other degree constructions. *Language and Linguistic Compass* 2.
- von Stechow, Arnim. 1984. Comparing theories of comparison. *Journal of Semantics* 3:1–77.
- Tomioka, Satoshi. 2007a. Intervention effects in focus: From a Japanese point of view. In *Isis working papers of the sfb 632*, ed. Shinichiro Ishihara, volume 9. Potsdam University.
- Tomioka, Satoshi. 2007b. Pragmatics of LF intervention effects: Japanese and Korean wh-interrogatives. *Journal of Pragmatics* 39:1570–1590.
- Yoshimura, Keiko. 2007. What does only assert and entail? In *Lodz papers in pragmatics*, volume 3, 97–117.
- Zeevat, Henk. 2009. “Only” as a mirative particle. In *Focus at the syntax-semantics interface: Working papers of the sfb 732*, ed. Arndt Riester & Edgar Onea, volume 3, 121–141. Stuttgart University.

The meaning of *what**

MARTINA WILTSCHKO
ICREA/Universitat Pompeu Fabra

1 Introduction

To initiate a conversation, we have to make sure we have the attention of our intended interlocutor — we have to summon them, as it were. There are various ways to achieve this goal, like for example calling them by their name, sometimes preceded by an attention getting particle (English *hey*). Within conversation analysis such initiating moves are referred to as SUMMONS, a term I adopt here (Schegloff 1968, 2007). In turn, the interlocutor has various ways to indicate that they are indeed paying attention and ready for interaction. For example, they could respond by uttering *what* (accompanied by falling intonation indicated by \Downarrow ¹). This is illustrated in (1), where \mathcal{I} stands for Initiator and \mathcal{R} stands for Responder.

- | | | | |
|-----|---------------|-------------------|---------|
| (1) | \mathcal{I} | Hey, Hotze! | SUMMONS |
| | \mathcal{R} | What \Downarrow | ANSWER |

The goal of this contribution is to explore the meaning of *what* when it functions as the answer to a summons move. I will show that it serves to indicate \mathcal{R} 's readiness for interaction by producing a response which simultaneously signals that \mathcal{R} is responding and that they want to know what \mathcal{I} wants. In this way, the utterance of a single wh-word *what* serves a complex function. Crucially, in Section 2, I demonstrate that this

* If I could have, I would have written a paper entitled “What is meaning?”. But I’m not there. So, I settled on a small case-study on the meaning of “what” (in context). And maybe that’s what you will appreciate more anyways, Hotze. I know you like to think about the details. So here is a little detail, with perhaps some larger big picture implications. And maybe on the occasion of your birthday, we can arrange for a meeting, to talk about this larger question of what meaning really is. And we can decide on the meaning of *what* we will talk about. The meaning of meaning ... the meaning of life ... the meaning of birthdays ... I hope the meaning of this particular one of yours is “Zufriedenheit”.

¹ Throughout this paper, I will mark punctuation in examples only when intonation is not marked via a downward or upward arrow. The presence of a question mark indicates the intended interpretation (i.e., question) but does not consistently correspond to rising intonation.

interpretation does not come about via ellipsis of a full-fledged wh-question (e.g., *what ~~do you want~~*). Instead, I argue that this complex meaning derives from its lexical meaning in combination with the grammatical structure that regulates linguistic interaction, namely the interactional spine (Wiltschko 2021), which I introduce in Section 3. In terms of its lexical meaning, *what* can be characterized as a variable restricted to inanimate individuals, which may include situations, propositions, or — as in the case of (1) — moves. In Section 4, I argue that *what* associates with a complex response structure which simultaneously marks the utterance as a response and requests a response from the interlocutor. This serves as a signal of \mathcal{R} 's readiness for interaction. Thus, I propose that the meaning of *what* is enriched with meaning that derives from the interactional spine. I conclude that this analysis makes for a more economic modelling of meaning than one that places the burden solely on the denotation of a lexical entry (Section 5).

2 Against an ellipsis analysis

When considering the question as to what *what* means when it functions to answer a summons, an obvious hypothesis to consider is that it stands in for a full wh-question via ellipsis. This is illustrated in (2), where *what* occupies the specifier of CP (as is typically the case for wh-words in English) and the remainder of the clause is elided (indicated by strike-through).

(2) An ellipsis analysis:

\mathcal{I} Hey, Hotze!
 \mathcal{R} [CP What [~~do you want~~]]

The ellipsis analysis in (2) is plausible for the following reasons. First, full wh-questions with initial *what* are possible as a response to a summon move, as shown in (3).

(3) \mathcal{I} Hey Hotze!
 \mathcal{R} a. **What** do you want?
b. **What** can I do for you?
c. **What**'s the matter?

Furthermore, ellipsis of this form is otherwise well-formed, as shown in (4) to (6). Here the initiation move is not a dedicated attention-getting

move (i.e., it may be uttered in the context of an ongoing conversation). The reaction move that follows can be either the single-word utterance *what* or a full-fledged question which repeats the question embedded in the initiating move.

- (4) \mathcal{I} You know what I want?
 \mathcal{R} a. **What** \Downarrow
 b. **What** do you want?
- (5) \mathcal{I} You know what you could do for me?
 \mathcal{R} a. **What** \Downarrow
 b. **What** can I do for you?
- (6) \mathcal{I} You know what's the matter?
 \mathcal{R} a. **What** \Downarrow
 b. **What's** the matter?

In these examples, the question in the initiating move serves as the antecedent for the elided string in the responses in (4) to (6). In contrast, it is not clear what might serve as the antecedent for the hypothetical ellipsis in (2): the initiating move consists of an attention getting particle combined with a vocative but there is no relevant propositional content.

What is even more striking is that the summons need not even be a verbal utterance, yet *what* is a possible response, as in (7).

- (7) \mathcal{I} stares at Hotze / taps Hotze on the shoulder / waves at Hotze
 \mathcal{R} **What** \Downarrow

In the summons in (7), there is no antecedent that would license an ellipsis in the answer. One might hypothesize that in these cases the antecedent is somehow implicit, in a way to be made precise. If so, the use of *what* when used as an answer to a summons would be akin to the cases in (4) to (6) where it precedes an elided clause.

There are, however, two problems which rule out the ellipsis analysis. First, consider the fact that *what* appears to be the only wh-word that can be used as a response to a summons. For example, *why* is not possible as a reaction to a vocative (8a) even though full *why* questions are, as shown in (8b,c).

- (8) \mathcal{S} Hey, Hotze!
 \mathcal{R} a. ***Why** \Downarrow
b. **Why** are you calling me?
c. **Why** do you need my attention?

Note that in the presence of a full antecedent, ellipsis is possible even following *why*, as shown in (9) and (10).

- (9) \mathcal{S} You know why I'm calling you?
 \mathcal{R} a. **Why** \Downarrow
b. **Why** are you calling me?
- (10) \mathcal{S} You know why I need your attention?
 \mathcal{R} a. **Why** \Downarrow
b. **Why** do you need my attention?

Given that both *what* and *why* questions are possible as answers to summons and given that both *what* and *why* questions license ellipsis, it is not clear why only *what* but not *why* can be used after a summons. This invites the conclusion that *what* in (1) is not an instance of an elided question.

This conclusion is supported by a further problem the ellipsis analysis faces: *what* is not possible in all situations even when *what* questions are. This is shown in (11), where there is no initiation move (and thus no summons). In this context bare *what* is ill-formed whereas a full *what* question is possible.

- (11) Upon entering a room where two people are fighting:
 \mathcal{S} a. ***What** \Downarrow
b. **What** is going on?

Note that there is nothing wrong with this particular wh-question such that it would not license ellipsis. It does when there is an appropriate antecedent in the initiating move, as in (12).

- (12) \mathcal{S} You know what's going on?
 \mathcal{R} a. **What** \Downarrow
b. **What** is going on?

A final piece of evidence against the ellipsis analysis of *what* in (1) is that it is necessarily associated with falling intonation. In this context, rising intonation is infelicitous, as shown in (13). This contrasts with full-fledged wh-questions, which can be realized with either falling or rising intonation, as in (13c,d).²

- (13) *S* Hey, Hotze!
R a. **What** ↘
 b. ***What** ↗
 c. **What** do you want ↘
 d. **What** do you want ↗

For completeness note that when *what* is clearly used with an elided clause, both rising and falling intonation are possible, as shown in (14).

- (14) *S* You know what's going on?
R a. **What** ↘
 b. **What** ↗
 c. **What** is going on ↘
 d. **What** is going on ↗

I conclude that when *what* is used as an answer to a summons it cannot simply be a wh-word followed by an elided clause. In the following sections I develop an alternative analysis. I argue that as an answer to a summons, *what* functions as a purely interactional unit of language. It is used to further the conversational interaction, without conveying propositional content.

3 Interactional structure as a source of interactional meaning

The main idea I wish to introduce here is that the interactional meaning of *what* can be understood as involving the working of an abstract system: the grammar of interactional language in the sense of Wiltschko (2021). That conversations are regulated by a system which is part of our competence is the hallmark of conversation analysis (Sacks et al. 1974). Wiltschko (2021) combines this insight of conversation analysis with those of generative grammar. With an in-depth investigation of units of language (UoLs) that contribute to the interaction itself, rather than to its

² See Hedberg et al. (2011) for an overview of the felicity conditions on rising and falling intonation in wh-questions.

content, Wiltschko (2021) concludes that the same system which configures the content of interaction also configures the logic of the interaction itself. The UoLs she explores are confirmational and response markers. The former define initiating moves and include utterance-final particles like *eh* and *huh*. The latter define reaction moves and include response markers such as utterance-initial *yeah* and *no*.

The core argument that there is a grammar of interactional language stems from the fact that the class of confirmational and the class of response markers display the same patterns of multi-functionality, and they do so across unrelated languages. For example, response markers like *yeah* can be used to answer questions, indicate agreement, acknowledge the move of an interlocutor, or simply mark a response as such. Wiltschko (2021) argues that this multi-functionality indicates the presence of an underlying abstract system, the so-called interactional spine, which enriches the interpretation of the UoLs themselves. That is, multi-functionality does not come about because of a series of homophonous UoLs, but instead because a given UoL associates with the spine in different positions and hence is enriched with different components of meaning as provided by the spine (see also Wiltschko 2014). The interactional spine consists of a grounding layer, responsible for the construction of common ground, and a response layer, responsible for the regulation of the interaction. Like all layers of structure on the spine, the response layer consists of a head position, which relates two arguments by asserting whether they coincide or not. The coincidence feature is an intrinsic property of every syntactic head and is valued by the UoLs that associate with it. The argument introduced by the response layer is the so-called *response set*, a set of elements that the interlocutors tend to (roughly corresponding to the *table* in the framework of Farkas and Bruce 2010). It can be indexed to the speaker or to the addressee, thus defining different move types. Initiating moves are defined by an addressee-oriented response set, as in (15a), while reaction moves are defined by a speaker-oriented response set, as in (15b).

- (15) a. INITIATION: [_{RespP} Resp-set_{Adr} [+/-coin] ...]
b. REACTION: [_{RespP} Resp-set_{Spkr} [+/-coin] ...]

In an initiating move, the content of the utterance is asserted to be or not to be in the addressee's response set. Thus, *RespP* allows a speaker to explicitly mark an utterance as requiring a response. Conversely, a reacting move can be marked as such by asserting whether or not the utterance is in the speaker's response set.

Furthermore, Wiltschko (2021) argues that RespP can be stacked, but only in a limited way. Specifically, a speaker-oriented RespP (i.e., a reaction move) can be embedded inside an addressee-oriented RespP (i.e., an initiation move). This configuration defines complex moves that simultaneously serve as a reaction and an initiation move, as shown in (16).

- (16) \mathcal{S} I got a new dog.
 $\mathcal{R}+\mathcal{S}$ Yeah ↗
 \mathcal{R} Yeah. I just got him last week. His name is Yoshi.

In (16), the response marker *yeah* is realized with rising intonation and it simultaneously serves as a reaction and an initiation. Specifically, with the use of *yeah*, the responder indicates that they are accepting the proposition but with the rising intonation, they indicate that further confirmation is required. This turns the reaction into an initiation (see Allwood et al. 1992). According to Wiltschko (2021), the complexity of the move correlates with the complexity of the interactional structure, as illustrated in (17): *yeah* associates with the speaker-indexed RespP marking a reaction while the rising intonation associates with the addressee-oriented RespP marking an initiation.³

- (17) [_{RespP} Resp-set_{Adr} [_{Resp} ↗ [+coin] [_{RespP} Resp-set_{Spkr} *yeah* [+coin] ...]]]

In what follows, I argue that these ingredients of the interactional spine allow for an analysis of the interactional use of *what* in (1).

4 *What* is interactional

I propose that the use of *what* in (1) is purely interactional. By this I mean that it is not used to inquire about any type of propositional content as is the case in typical content questions (e.g., *What do you want?*). Instead, it is used to inquire about the nature of the interaction. I argue that this is a result of associating *what* with the interactional structure. Specifically, I propose that *what* associates with a complex RespP of the type illustrated in (17): it simultaneously functions as a reaction and an initiation move. Specifically, *what* is intrinsically pronominal and is interpreted as an indefinite variable, restricted to inanimate entities,

³An addressee-oriented RespP cannot further be dominated: once the utterance is put into the interlocutor's response-set, the current speaker has to end their turn.

including propositions and the like. Due to its syntactic position within the specifier position of RespP, it is interpreted as an interactional variable. This aspect of the analysis is in line with Krifka's (2013) analysis of propositional anaphors. Specifically, Krifka argues that *that* can be used to anaphorically relate to propositions or speech acts, as shown in (18).

- (18) \mathcal{I} Ede stole the cookie.
 \mathcal{R}_1 I didn't know **that**. (antecedent = proposition)
 \mathcal{R}_2 **That's** a lie! (antecedent = speech act)
 (adapted from Krifka 2013:4, ex. 19)

Note that *that* and *what* differ in terms of definiteness: *that* functions as a (definite) pronominal whereas *what* functions as an indefinite variable and hence can serve as a question word. In fact, in many languages wh-words are interpreted as indefinite pronouns when they remain in situ. This is shown in (19) for Dutch.

- (19) a. **Wat** heb je gedaan? b. Jan heeft **wat** gedaan.
 what have you done John has what done
 'What have you done?' 'John has done something.'
 (Postma 1994:187, ex. 2)

I hypothesize that the interpretation of *what* in the interactional structure proceeds as follows. In the speaker-oriented RespP, the use of the indefinite indicates that there is an initiation move which serves as the trigger for the current reaction move. However, the content of this initiation is unknown. In the addressee-oriented RespP, the use of the indefinite indicates that the speaker requests a response. This is schematized in (20).

- (20) [_{RespP} Resp-set_{Adr} *what* [_{RespP} Resp-set_{Spkr} ~~*what*~~ ...]]

Thus, I hypothesize that by using *what* as an interactional but indefinite variable the speaker signals that they are ready for interaction, i.e., that they accept the summons. Note that the interactional spine plays a key role in this analysis: the fact that *what* in this context is interpreted as an interactional variable is syntactically, rather than lexically, conditioned.

In what follows, I show how the analysis in (20) accounts for the properties of interactional *what*, which pose a problem for the ellipsis

analysis. First, we have seen that interactional *what* is well-formed even if there is no propositional antecedent, such as after a bare vocative or even a non-verbal event like a stare or a shoulder tap. While ellipsis requires a verbal antecedent, pronominal forms do not, as shown in (21).

(21) Upon hearing an unexpected noise:

ℳ What was **that**?

A second property of interactional *what*, which cannot be accounted for with an ellipsis analysis, has to do with the fact that it requires an explicit initiation move to be well-formed (as illustrated in (11)). The analysis in (20) captures this: *what* in RespP is interpreted as a pro-form for a RespP, which by definition must be an interactional move.

The third property of interactional *what* is that it has to be realized with falling intonation while rising intonation is ill-formed. This differs from regular wh-questions including those that contain elided material (see the contrast between (13) and (14)). I propose that this restriction is also syntactically conditioned. Specifically, according to Wiltschko (2021) rising intonation is associated with RespP_{Adr} and indicates the request for a response. Hence interactional *what* in (20) is in complementary distribution with rising intonation. In contrast, in the context of a regular wh-question, rising *what* does not associate with RespP (but instead is located in SpecCP). Thus, rising intonation can associate with RespP. As for falling intonation, I assume that it is not interpreted as a (meaningful) intonational tune and hence is not associated with the interactional spine (Wiltschko 2022). Instead, falling intonation derives from the absence of a marked intonation. Since pitch declines automatically with the decrease in subglottal air pressure (Cohen et al. 1982), absence of a marked intonational tune is realized as falling intonation.

Finally, the last property of interactional *what*, which sets it apart from propositional *what* and hence cannot be accounted for with an ellipsis analysis, has to do with the fact that it is restricted to *what*. As we have seen, *why* is not possible even though full propositional *why* questions are perfectly sensible in similar contexts (see (8) to (10)). The analytical challenge for the interactional analysis of *what* boils down to the question as to why *why* cannot be associated with the interactional spine to react to a summons. I tentatively propose that this has to do with the presuppositions associated with *why*. To see this, consider regular content questions. A *what*-question can be responded to by denying that

there is something that corresponds to the variable introduced by *what*. This is shown in (22) and (23).

(22) *S* **What** did you eat?
R **Nothing.**

(23) *S* **What** do you want?
R **Nothing.**

A *why* question on the other hand presupposes that the event whose reason is being questioned has happened. This is shown in (24) and (25).

(24) *S* **Why** did you eat?
R #I didn't eat.

(25) *S* **Why** are you calling me?
R #I'm not calling you.

Crucially, an attention getting move is a special kind of initiation, which may occur simply to attract the attention of the interlocutor but without conveying content (Filipi 2009). Since *what* does not presuppose content, it is compatible with this use. In contrast, when using *why* the responder has to be sure that there is in fact an initiation and that this is shared knowledge.

5 Conclusion

The goal of this contribution was to explore the meaning of *what* when it is used to react to a summons. We have seen that despite the apparent simplicity of the move its function is complex: it serves as a reaction to the summons as well as a request for a response. As such it indicates readiness for interaction. I have demonstrated that this cannot be derived from an ellipsis analysis which would attribute this complexity to an elided clause. Rather, I proposed that *what* in this context associates with the interactional spine. Its intrinsic meaning (an indefinite variable) is enriched with the meaning that comes with the interactional spine. Specifically, it associates with both the speaker-oriented and the addressee-oriented RespP and thus simultaneously marks a reaction and initiation move. I submit that this analysis is more economical than one that would postulate a dedicated lexical entry for this specific use as the

interactional spine has been motivated on independent grounds (Wiltschko 2021). More generally, the exploration of interactional *what* speaks to the importance of combining insights from conversation analysis and formal grammatical analysis. While expressive and use-conditional aspects of meaning play an increasing role within semantic theory (Potts 2007; Gutzmann 2013, 2015), which was traditionally concerned with truth-conditional meaning, the contribution of conversational interaction has, to date, received less attention (but see Ginzburg 2012). I hope to have shown that interactional meaning should be integrated into our notion of meaning. It provides a rich empirical domain and presents novel challenges regarding its integration into formal theories of semantics and pragmatics.

References

- Allwood, Jens, Joakim Nivre, & Elisabeth Ahlsén. 1992. On the semantics and pragmatics of linguistic feedback. *Journal of Semantics* 9(1):1–26.
- Cohen, Antonie, René Collier, & Johan ‘t Hart. 1982. Declination: Construct or intrinsic feature of speech pitch? *Phonetica* 39(4–5):254–273.
- Farkas, Donka F. & Kim B. Bruce. 2010. On reacting to assertions and polar questions. *Journal of Semantics* 27(1):81–118.
- Filipi, Anna. 2009. *Toddler and Parent Interaction: The Organization of Gaze, Pointing and Vocalization*. Amsterdam, Netherlands: John Benjamins.
- Ginzburg, Jonathan. 2012. *The Interactive Stance*. Oxford, UK: Oxford University Press.
- Gutzmann, Daniel. 2013. Expressives and beyond. An introduction to varieties of use-conditional meaning. Daniel Gutzmann & Hans-Martin Gärtner (eds.), *Beyond Expressives. Explorations in Use-Conditional Meaning*. Leiden, Netherlands: Brill: 1–58.
- Gutzmann, Daniel. 2015. *Use-conditional Meaning: Studies in Multidimensional Semantics*. Oxford, UK: Oxford University Press.
- Hedberg, Nancy, Juan M. Sosa, & Emrah Görgülü. 2011. On the meaning of non-canonical question intonation. *Experimental and Theoretical Advances in Prosody* 2:23–25.

- Krifka, Manfred. 2013. Response particles as propositional anaphors. *Proceedings of Semantics and Linguistic Theory (SALT)* 23:1–18.
- Postma, Gertjan. 1994. The indefinite reading of WH. *Linguistics in the Netherlands* 11(1):187–198.
- Potts, Christopher. 2007. The expressive dimension. *Theoretical linguistics* 33.2: 165–198.
- Sacks, Harvey, Emanuel Schegloff, & Gail Jefferson. 1974. A simplest systematics for the organization of turn-taking for conversation. *Language* 50:696–735.
- Schegloff, Emanuel. 1968. Sequencing in conversational openings. *American Anthropologist* 70(6):1075–1095.
- Schegloff, Emanuel. 2007. *Sequence Organization in Interaction: A Primer in Conversation Analysis*. Cambridge, UK: Cambridge University Press.
- Wiltschko, Martina. 2014. *The Universal Structure of Categories. Towards a Formal Typology*. Cambridge, UK: Cambridge University Press.
- Wiltschko, Martina. 2021. *The Grammar of Interactional Language*. Cambridge, UK: Cambridge University Press.
- Wiltschko, Martina. 2022. What is the syntax-pragmatics interface. In Gabriela Bilbiie, Berthold Crysmann, & Gerhard Schaden (eds.), *Empirical Issues in Syntax and Semantics* 14. Paris, France: CSSP, 273–308. URL: <http://www.cssp.cnrs.fr/eiss14/>

Modal adverbs, conjunction reduction, and the structure of coordination*

JAN-WOUTER ZWART
University of Groningen

1 Introduction

The earliest generative analyses of coordination (Chomsky 1957:113; Gleitman 1965:273–274) took noun phrase coordination (1a) to be derived from clausal coordination via conjunction reduction (CR) (1b):

- (1) a. John and Mary went to the store.
b. John ~~went to the store~~ and Mary went to the store.

This derivation is problematic in the case of so-called symmetric predicates (or non-Boolean coordination, cf. Schmitt 2021), where the distributive reading inherent to the underlying clausal conjunction is not available (*pace* Schein 2017):

- (2) a. John and Mary went to the store together.
b. *John went to the store together and Mary went to the store together.

This led Lakoff and Peters (1969:114) to propose an additional mechanism of straightforward noun phrase coordination (also Jackendoff 1977:51). But as Schmitt (2021:14) notes, this is unsatisfactory in view

* I first met Hotze Rullmann when I transferred from Nijmegen to the University of Groningen linguistics program as an undergraduate student in the Fall of 1985. Hotze was already enrolled there, and we took classes together from then on. In fact, we graduated on the very same day, July 29, 1988, and, on Hotze's suggestion, had a joint party afterwards in one of the cafés in Kleine Kromme Elleboog. Soon after that, Hotze left for Amherst, and we pursued our PhDs separately. I still cherish the copy of Hotze's MA-thesis 'Referential dependency', which I thought was brilliant. It turned out to be a reliable indicator of what was to come. After a couple of years, Hotze rejoined the faculty in Groningen in the context of Jack Hoeksema's project on negation, and we had great fun collaborating during that period. Afterwards it was immensely satisfying to see how Hotze thrived in Canada and it was always a pleasure to meet up, here in Groningen, or wherever. Some of the examples in this paper were inspired by an amusing discussion we had at some point about the final seconds in college basketball games. As Hotze knows, these final seconds can last forever — fortunately, of course, this is also true of the post-Festschrift time in Academia.

of the fact that languages invariably use the same conjunction for the two types of coordination (sentential *cum* CR and phrasal *sans* CR).

Moreover, as shown by Dougherty (1970/1971) and Lasersohn (1995), a sentential coordination *cum* CR analysis can be dispensed with for coordination with non-symmetric predicates as well. What needs to be explained, in that case, is the availability of a distributive reading (3b), in which John and Mary went to different stores or to the same store on different occasions:

- (3) John and Mary went to the store *can mean*
- a. John and Mary went to the store together. (collective)
 - b. John went to the store and Mary went to the store.
(distributive)

As Lasersohn (1995) argues, the distributive reading of (3) can be analysed as residing in the semantics of the verb phrase. This is because predicates with plural (or conjoined) subjects can be taken to ambiguously predicate over groups or the individuals making up those groups (Lasersohn 1995:85). With symmetric predicates, only predication over groups is available (Lasersohn 1995:86). The distributive reading, then, can be accounted for without resorting to underlying sentential coordination.

The feasibility of a CR-analysis of coordination comes up in the analysis of conjunctions featuring a modal adverb, as discussed in Collins (1988) (hence called ‘Collins conjunctions’):

- (4) John and perhaps Mary went to the store.

Since *perhaps* is a clausal modifier, an analysis involving reduction of underlying clausal coordination suggests itself (though Collins 1988:17f rejects it — rightly, as we will see). In this connection, it may be seen as relevant that (4) disallows symmetric predicates:

- (5) *John and perhaps Mary went to the store together.
(cf. *John went to the store together and perhaps Mary went to the store together.)

On the assumption that (4) is derived via CR, the ungrammaticality of (5) could be explained through the absence of an underlying clausal

coordination source. But a CR-analysis is called into question again by cases like (6):

- (6) John, Bill and perhaps Mary went to the store together.

In this article I adduce additional arguments against a CR-analysis of examples like (4), supporting Collins (1988), though not necessarily his conclusion that *perhaps* modifies the conjunction *and* (Collins 1988:12). The analysis leads to a better understanding of the distribution of modal adverbs (Section 3), and casts doubt on the common conception of coordinate structures as being headed by the conjunction (Section 4).

2 Against conjunction reduction with Collins conjunctions

The following arguments intend to show that a conjunction reduction (CR) analysis of sentences like (4) is not just unnecessary, but in fact impossible, or at least very problematic. We will switch back and forth between English and Dutch, guided by the question of which language allows for a better exposition of the problem.

2.1 Word order problems

Consider the example in (7) of a Collins conjunction case from Dutch.

- (7) Jan en misschien Marie kom-en ook.
 John and perhaps Mary come-PL also
 ‘John and perhaps Mary will also be there.’

If (7) were to be derived via clausal coordination and CR, the underlying structure would be (8). (The argumentation abstracts away from the verb morphology, which we assume is adjusted postsyntactically.)

- (8) Jan kom-t ook en Marie kom-t misschien ook.
 John come-3SG also and Mary come-3SG perhaps also

But applying CR to (8) yields (9), not (7).

- (9) Jan ~~kom-t~~ ~~ook~~ en Marie kom-en misschien ook.
 John come-3SG also and Mary come-PL perhaps also

In (9), *misschien* ‘perhaps’ is in the wrong position. To make sure *misschien* precedes *Marie*, *misschien* would have to be fronted. But Dutch being a ‘verb second’ language, fronting of *misschien* would also trigger verb movement, yielding (10), also unsuitable.

- (10) *Jan ~~k_{om}t~~—~~ook~~ en misschien kom-en Marie ook.
 John come-3SG also and perhaps come-PL Mary also

One might suppose (with Chomsky 2000:37) that verb movement, like deletion, is a postsyntactic process, and that it takes place after CR. This would lead us to think that CR gives us the intermediate stage in (11), with verb placement and agreement adjusted subsequently to yield (7).

- (11) *Jan ~~ook~~—~~k_{om}t~~ en misschien Marie ook kom-en.
 John also come-3SG and perhaps Mary also come-PL

There are at least two problems with that idea.

First, verb second places the finite verb to the right of the first constituent (*misschien* in (11)), so even if verb placement is a postsyntactic operation, it would turn (11) into (10), not (7).

Second, the analysis would have to assume that verb movement follows ellipsis (now both considered to be postsyntactic operations), but this is not generally the case. Right Node Raising, for instance, can only take place when the verb has vacated its original clause-final position, as in (12a):

- (12) a. Jan kook-t ~~de~~—~~groente~~ en Marie stoom-t
 John boil-3SG the vegetables and Mary steam-3SG
 de groente.
 the vegetables
 ‘John boils and Mary steams the vegetables.’

- b. *... dat Jan ~~de~~—~~groente~~ kook-t en Marie
 COMP John the vegetables boil-3SG and Mary
 de groente stoom-t.
 the vegetables steam-3SG

Intended: ‘... that John boils and Mary steams the vegetables.’

Verb movement, then, may be postsyntactic, but it cannot be strictly ordered after ellipsis.

2.2 No clausal source

Dutch has a number of exceptive particles (*behalve* ‘except’, *uitgezonderd* ‘with the exception of’) acting like coordinating conjunctions (Klein 1985; Komen 1993):

- (13) Iedereen behalve Jan kwam naar het feest.
 everybody except John come:PAST.SG to the party
 ‘Everyone except John came to the party.’

As noted by Van der Heijden (1999:128), these exceptive expressions lack a clausal source on which CR could operate to yield sentences like (13).

- (14) * Iedereen kwam naar het feest behalve Jan kwam naar het feest.

The correct expression would involve an embedded clause (which works with *behalve* but not with other expressions like *uitgezonderd*).

- (15) Behalve dat Jan naar het feest kwam.
 except COMP John to the party come:PAST.SG
 ‘Except that John came to the party.’

Interestingly, these exceptive conjunctions can be modified by *perhaps*:

- (16) Iedereen behalve misschien Jan kwam naar het
 everybody except perhaps John come:PAST.SG to the
 feest.
 party
 ‘Everyone except perhaps John came to the party.’

This shows that expressions like *perhaps* can modify conjunctions in the absence of a potential underlying clausal source.

Other elements modifying conjunctions also lack a clausal source for CR to operate on, either because the underlying clause would have to be an embedded clause, as with *behalve*, or because an underlying source is unavailable, as with *uitgezonderd*. The first category includes expressions like *ik sluit niet uit* ‘I won’t exclude [sc. the possibility]’ and *alles wijst erop* ‘all indications are’, the second includes expressions like *met een kleine slag om de arm* ‘with some hedging’ and *wat helemaal leuk zou zijn* ‘what would be especially good’. Space prevents me from

illustrating these here, but the problem for a CR-analysis of noun phrase coordination is clearly not an isolated case.

2.3 Obligatory narrow scope

Many observations indicate that *perhaps* in (4) has a narrower scope than would be expected if it were to modify an underlying clause in a CR-analysis. Consider the following real-life example.

- (17) This twelve point lead and perhaps victory for Ohio State is going to be in large part due to the play of Cotie McMahon.
(Rebecca Lobo, ESPN, March 25, 2023 with 1:10 left to play in the NCAA Women's Basketball Championship Sweet Sixteen matchup between Ohio State and Connecticut)

Here *perhaps* narrowly modifies *victory*, in keeping with commentators' tradition of allowing for spectacular turns of events with just seconds left in the game (as noted by Rullmann, p.c.). On a CR-derivation, *perhaps* would wrongly modify the attribution of the victory to McMahon's play:

- (18) #This victory for Ohio State is perhaps going to be in large part due to the play of Cotie McMahon.

Here's another example, from Dutch:

- (19) Jan en misschien Marie krijg-en zeker een beurs.
John and perhaps Mary get-PL certainly a grant
'John and perhaps Mary will certainly get a grant.'

This sentence can be used to describe a situation where there are various stages in the selection process for a particular grant, and it is certain that Jan has reached the final, decisive stage, where applicants are certain to obtain their grant, but it is not yet certain, though by no means excluded, that Mary will reach that stage. The source sentence for the CR-derivation would be the anomalous (20).

- (20) #Misschien krijg-t Marie zeker een beurs.
perhaps get-3SG Mary certainly a grant
'Mary will perhaps certainly get a grant.'

In (20), *misschien* ‘perhaps’ and *zeker* ‘certainly’ provide conflicting modifications of the clausal predication ‘Mary gets a grant’. This conflict is absent in (19), indicating that *misschien* in (19) does not have sentential scope. Rather, what (19) conveys is that Mary is perhaps a member of the group that will certainly get a grant (see Section 3).

A similar scope discrepancy is apparent with hendiadys coordinations like (21), which can be modified by a conjunction adverb.

- (21) a. My friend and future colleague will also attend.
 b. My friend and perhaps future colleague will also attend.

In (21b), it is possible to think that the person indicated by the hendiadys coordination will at some point in the future be my colleague. But the underlying clause in a CR-analysis (22) would have a different reading, namely that it is possible to think that a specific future colleague of mine will also attend.

- (22) #My friend will also attend and perhaps my future colleague will also attend.

So the hedging brought on by *perhaps* has a narrower scope than the CR-analysis would allow for.

Consider also the interaction of *perhaps* with modal verbs and negation. Doherty (1987:52) observes that the modal element of *perhaps* (‘it is possible to think’) in sentences like (23) has scope over negation:

- (23) Alice has perhaps not won. $\diamond > \neg$
 (i.e., It is possible to think that it is not the case [that Alice has won].)

When *perhaps* is not present, negation takes scope over the modal element expressed by verbs like *can*:

- (24) Alice cannot succeed. $\neg > \diamond$
 (i.e., It is not the case that it is possible [for Alice to succeed].)

When *perhaps*, modal verbs, and negation are combined, negation appears to take intermediate scope between the modal element of *perhaps* and the modal element of *can*:

- (25) Alice perhaps cannot succeed. $\diamond > \neg > \diamond$
(i.e., It is possible to think that it is not the case that it is possible
[for Alice to succeed].)

But now consider (26):

- (26) Baylor and perhaps UConn cannot advance to the Elite Eight.
 $\neg > \diamond$

This sentence can be used in a situation where Baylor has already been eliminated, and UConn is trailing in the match in which to qualify for the Elite Eight. As in (24), *not* takes scope over the modal *can*:

- (27) It is not the case that it is possible [for Baylor and perhaps UConn to advance].

However, the CR-source sentences for the derivation of (26) would include (28), where *perhaps* takes scope over *not*.

- (28) UConn perhaps cannot advance to the Elite Eight $\diamond > \neg > \diamond$
(i.e., It is possible to think that it is not the case that it is possible
[for UConn to advance].)

Again, *perhaps* when used as a conjunction adverb has a narrower scope than when used as a sentence adverb.

3 The status of adverbs in Collins conjunctions

The arguments in the previous section against a conjunction reduction analysis of coordinate structures like *John and perhaps Mary* in (4) call for an alternative analysis of the ungrammaticality of (5). Here we may follow Lasersohn's (1995:85) analysis of predicates with conjoined subjects as predicating over groups or the individuals making up the group — only the former being relevant in the case of symmetric predicates. The contribution of *perhaps* in (4) now appears to be that it renders the composition of the group uncertain: it can be either John or John and Mary (cf. Collins 1988:13). Example (5) then tells us that symmetric predicates must be compatible with either scenario.

In (6), there is a similar uncertainty about group composition, but here the two options are (i) John and Bill and (ii) John and Bill and Mary;

symmetric predication is compatible with both scenarios, hence (6) is not ungrammatical.

Interestingly, the effect of *perhaps* on group composition is the same with disjunctive subjects as in (29):

(29) John or perhaps Mary went to the store.

Disjunction differs from conjunction in that with conjunction, the predicate applies to the entire group (the composition of which may be in question in the case of Collins conjunctions), whereas with disjunction, the predicate applies to only a subset of the group. But this does not affect the group composition, and *perhaps* qualifies the possibility of Mary being part of the group in the same way with conjunctive and disjunctive coordination.

If this is on the right track, we may not have to conclude, as Collins (1988:12) does, that *perhaps* in (4) modifies the conjunction *and*. Rather, what seems to be the case is that *perhaps* modifies the inclusion of (the denotation of) Mary in the (denotation of the) coordinate noun phrase. The process of set inclusion referred to here is perhaps sufficiently similar to the process of set inclusion that joins the (denotation of the) subject and the (denotation of the) predicate in standard subject-predicate combination (cf. Lasersohn 1995:85). If it is this process of set inclusion that makes the addition of a modal adverb possible, we can explain the appearance of adverbials in coordinate structures without having to resort to a questionable derivation involving conjunction reduction, or to an analysis in which the adverb is adjoined to the conjunction itself, as in Collins (1988:§4).

At the same time, we can now also understand why *perhaps* cannot normally (i.e., outside the context of coordinate structures) modify noun phrases:

(30) Perhaps Mary went to the store.

In (30), *perhaps* can only be interpreted as modifying the clausal predication — there is no narrow reading possible in which *perhaps* modifies only *Mary*. But this is because outside the context of coordinate structures, noun phrases do not give rise to a process of set inclusion that adverbs like *perhaps* can qualify.

If this is the correct analysis of modal adverbs in coordinate structures, we predict that all propositional modifiers, such as negation markers, adverbs like *probably*, evidential markers like *I hear*, and focus

markers like *also* and *only*, should be able to appear inside coordinate structures. This prediction is confirmed:

- (31) a. John and not Mary went to the store.
b. John and probably Mary went to the store.
c. John and I hear Mary went to the store.
d. John and also/only Mary went to the store.

In all these cases, the element preceding *Mary* modifies the inclusion of (the denotation of) *Mary* in (the denotation of) the subject group, underscoring the proposed similarity of coordination and predication.

4 Conclusion: the structure of coordination

Collins (1988) discusses the nature and position of modal adverbs in coordinate structures in the context of an analysis of the phrase structure of coordinations. The idea that *perhaps* in (4) modifies the conjunction *and* leads him to conclude that *and* is the head of the coordinate structure, and that the coordinate structure as a whole has the standard X'-structure in (32) (see Progovac 2003:260ff for further discussion).

- (32) [_{ConjP} John [_{Conj'} [_{Conj} and] Mary]]

The ConjP analysis is motivated empirically by the observation of asymmetries between the two conjuncts (Johannessen 1993; Haspelmath 2007:9), and conceptually by the conformity it brings with the general X'-theory of phrase structure (e.g., Kayne 1994:57). In light of minimalist thinking about the derivation of phrase structure, neither argument is compelling (see Zwart, to appear, for more discussion).

Crucially, the only structure building operation available in the minimalist model of grammar is Merge, which creates a pair of sisters. This is in fact an asymmetric pair, either automatically, because that is the way Merge operates (Zwart 2009:163), or indirectly, because of a labeling algorithm turning an unordered set into an ordered pair (Chomsky 2000:133). Either way, the simplest structure is a pair of sisters, and the asymmetries between the conjuncts that have been observed in the literature may be accounted for if the sister pair is asymmetric. Any further elaboration upon that simplest (headless) structure would have to be explicitly motivated.

If I am correct in this article, the distribution of modal adverbs like *perhaps* inside coordinate noun phrases should not be interpreted to yield

the required explicit motivation for the head status of the conjunction. This leaves open the possibility that coordinated phrases are just pairs of sisters, the second of which may be marked by a coordinating element in a variety of ways, and may be modified by a modal adverb to qualify the inclusion in the group denoted by the coordinate structure.

References

- Chomsky, Noam. 1957. *Syntactic Structures*. The Hague, Netherlands: Mouton.
- Chomsky, Noam. 2000. Minimalist inquiries: the framework. In Roger Martin, David Michaels, & Juan Uriagereka (eds.), *Step by step: essays on minimalist syntax in honor of Howard Lasnik* Cambridge, MA: MIT Press, 89–155.
- Collins, Chris. 1988. Conjunction adverbs. Unpublished manuscript, Massachusetts Institute of Technology.
- Doherty, M. 1987. Perhaps. *Folia Linguistica* 21:45–65.
- Dougherty, Ray C. 1970/1971. A grammar of coordinate conjoined structures. *Language* 46:850–898 and 47:298–339.
- Gleitman, Lila R. 1965. Coordinating conjunctions in English. *Language* 41: 260–293.
- Haspelmath, Martin. 2007. Coordination. In Timothy Shopen (ed.), *Language typology and syntactic description, volume 2: complex constructions*. Cambridge, UK: Cambridge University Press, 1–51.
- Jackendoff, Ray. 1977. *X' Syntax: A Study of Phrase Structure*. Cambridge, MA: MIT Press.
- Johannessen, Janne B. 1993. *Coordination: A Minimalist Approach*. Ph.D. dissertation, University of Oslo.
- Kayne, Richard S. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.
- Klein, Maarten. 1985. *Behalve: voorzetsel of voegwoord?* *De Nieuwe Taalgids* 78:363–368.
- Komen, J. A. M. 1993. De uitzonderlijkheid van *uitgezonderd*. *De Nieuwe Taalgids* 86:127–148.

- Lakoff, George & Stanley Peters. 1969. Phrasal conjunction and symmetric predicates. In David A. Reibel & Sanford A. Schane, (eds.), *Modern Studies in English: Readings in Transformational Grammar*. Englewood Cliffs, NJ:Prentice-Hall, 113–142.
- Lasersohn, Peter. 1995. *Plurality, Conjunction and Events*. Dordrecht, Netherlands: Kluwer.
- Progovac, Ljiljana. 2003. Structure for coordination. In Lisa Cheng & Rint Sybesma (eds.), *The second Glot International State-of-the-Article book: the latest in linguistics*. New York, NY: Mouton de Gruyter, 241–287.
- Schein, Barry. 2017. *'And': Conjunction Reduction Redux*. Cambridge, MA: MIT Press.
- Schmitt, Viola. 2021. Boolean and non-Boolean conjunction. In Daniel Gutzmann, Lisa Matthewson, Cécile Meier, Hotze Rullmann, & Thomas Ede Zimmermann (eds.), *The Wiley Blackwell Companion to semantics*, chapter 111. Hoboken, NJ: John Wiley and Sons, Ltd.
- Van der Heijden, Emmeken. 1999. *Tussen nevenschikking en onderschikking: een onderzoek naar verschillende vormen van verbinding in het Nederlands*. Ph.D. dissertation, Catholic University Nijmegen.
- Zwart, Jan-Wouter. 2009 Prospects for top-down derivation. *Catalan Journal of Linguistics* 8:161–187.
- Zwart, Jan-Wouter. (to appear). Coordination. In Sjef Barbiers, Norbert Corver, & Maria Polinsky (eds.), *The Cambridge Handbook of Comparative Syntax*. Cambridge, UK: Cambridge University Press.