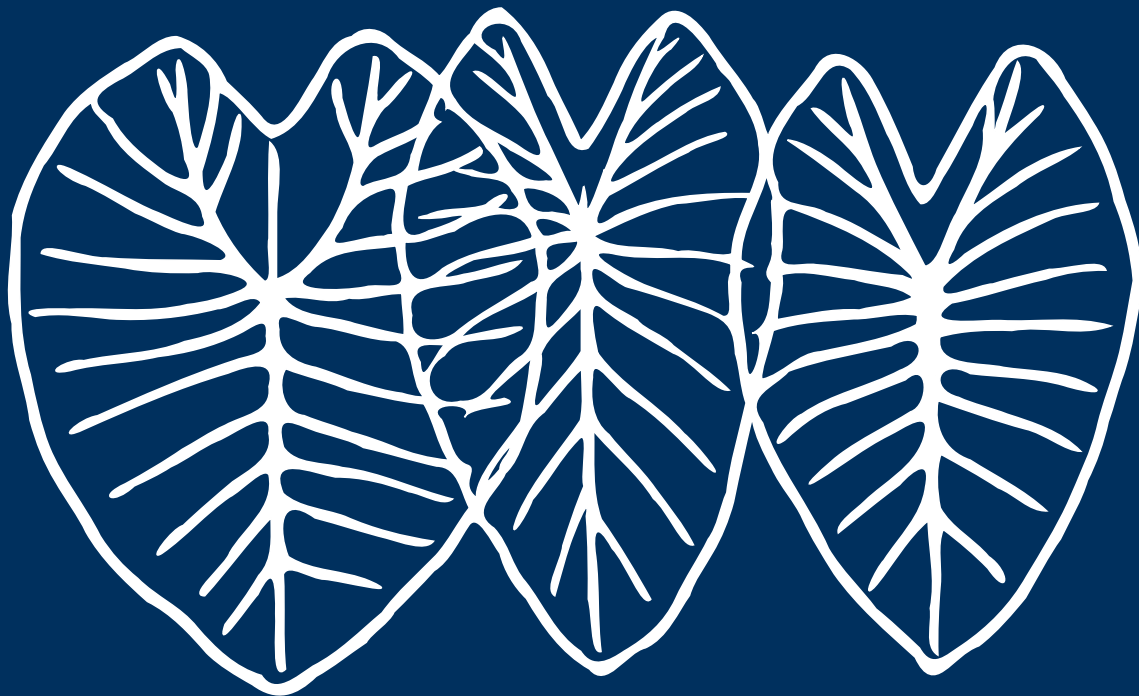


Proceedings of TripleA 10

Fieldwork Perspectives on the Semantics of African, Asian and Austronesian Languages



Edited by Jeanne Lecavelier, Niklas Geick, Mira Grubic,
Prarthanaa Bharadwaj and Malte Zimmermann

This work is protected by copyright and/or related rights. You are free to use this work in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you need to obtain permission from the rights-holder(s).

<https://rightsstatements.org/page/InC/1.0/?language=en>

Proceedings of TripleA 10:

Fieldwork Perspectives on the Semantics of African, Asian and Austronesian Languages

Edited by Jeanne Lecavelier, Niklas Geick, Mira Grubic, Prarthanaa Bharadwaj and Malte Zimmermann.
2024

Cover design by Vera Hohaus and Jeanne Lecavelier.

Taro illustrations from Leo D. Whitney, F. A. I. Bowers & M. Takahashi (1939), "Taro Varieties in Hawaii", Hawaii Agricultural Experiment Station Bulletin 84, Fig. 2, p. 16a-b.

<https://scholarspace.manoa.hawaii.edu/handle/10125/4327>

Published online on the Publication Server of the University of Potsdam:

<https://doi.org/10.25932/publishup-64798>

<https://nbn-resolving.org/urn:nbn:de:kobv:517-opus4-647980>

Preface

The TripleA workshop series was founded in 2014 by linguists from Potsdam and Tübingen with the aim of providing a platform for researchers that conduct theoretically-informed linguistic fieldwork on meaning. Its focus is particularly on languages that are under-represented in the current research landscape, including but not limited to languages of Africa, Asia, and Australia, hence TripleA.

For its 10th anniversary, TripleA returned to the University of Potsdam on the 7-9th of June 2023.

The programme included 21 talks dealing with no less than 22 different languages, including three invited talks given by Sihwei Chen (Academia Sinica), Jérémy Pasquereau (Laboratoire de Linguistique de Nantes, CNRS) and Agata Renans (Ruhr-Universität Bochum). Nine of these (invited or peer-reviewed) talks are featured in this volume.

We are grateful to the following reviewers for their time and feedback on the abstracts:

Polina Berezovskaya	Vera Hohaus	Roland Pfau
Rajesh Bhatt	Jens Hopperdietzel	Josh Phillips
M. Ryan Bochnak	Andrew Koontz-Garboden	Agata Renans
Kenyon Branam	Natasha Korotkova	Viola Schmitt
Sihwei Chen	Ana Krajinovic	Hooi Ling Soh
Victoria Chen	Lisa Matthewson	Sandhya Sundaresan
Elizabeth Coppock	Jon Ander Mendia	Guillaume Thomas
Christopher Davis	Paula Menendez-Benito	Jenneke van der Wal
Virginia Dawson	Anne Mucha	Jozina Vanderklok
Reginald Akuoko Duah	Keely New	Kilu von Prince
Michael Yoshitaka Erlewine	Augustina Pokua Owusu	
Martina Faller	Jérémy Pasquereau	
John Gluckman	Matt Pearson	

TripleA 10 was funded by the Department of Linguistics of the University of Potsdam and the Project C02 of the Collaborative Research Center SFB 1287 "Limits of Variability in Language".

Organizing Committee:

Giuliano Armenante, Nadine Bade, Mira Grubic,
Jeanne Lecavelier, Mareike Philipp, Malte Zimmermann.

Helping team:

Zaher Alkaei, Prarthanaa Bharadwaj,
Niklas Geick, Ines Mauer.

Table of contents

Anne Mucha, James J. Engels, Fred Whibley and Wataru Uegaki (University of Edinburgh)	
<i>Negative modality in Hausa, Thai and Kîîtharaka</i>	1
James C. Wamsley (Indiana University)	
<i>Two Types of Definiteness: A Case Study from Hakha Lai</i>	18
Virginia Dawson (Western Washington University)	
<i>Inclusive plural in a ‘general number’ language</i>	33
Anastasija Gruzdeva (Institute of Linguistics, RAS / Lomonosov MSU), Anna Alhazova, Anna Golovnina, Regina Nasyrova (Lomonosov MSU) and Feodor Sadkovsky (Institute of Linguistics, RAS / Lomonosov MSU)	
<i>Kumyk Verb Classification: Event and Argument Structure</i>	47
Siena Weingartz and Vera Hohaus (University of Manchester)	
<i>Variable Modal Strength in Afrikaans and Samoan: Deriving Strong Necessity from Weak Necessity</i>	60
Ousmane Cisse and Elizabeth Coppock (Boston University)	
<i>Reduplicated Distributivity in Mandinka</i>	75
Badiba Olivier Agodio (Guébie Community Member), Peter Jenks, Hannah Sande (UC Berkeley) and Malte Zimmermann (Universität Potsdam)	
<i>Indexed definiteness without demonstratives in Guébie</i>	90
Polina Berezovskaya (University of Tübingen)	
<i>Variation in the Grammar of Alternatives — Are there Intervention Effects in Tundra Nenets?</i>	106
Sihwei Chen (Academia Sinica)	
<i>Towards the semantics of Atayal polar question particles and a semantic typology</i>	123
Agata Renans (Ruhr-Universität Bochum)	
<i>In search of exclusive plural — insights from Hausa</i>	142

Negative modality in Hausa, Thai and Kîtharaka¹

Anne Mucha — The University of Edinburgh

James J. Engels — The University of Edinburgh

Fred Whibley — The University of Edinburgh

Wataru Uegaki — The University of Edinburgh

Abstract. This paper presents, to our knowledge, the first attempt to describe crosslinguistic variation in the lexicalization of impossibility modals. We present novel data on the expression of negative modality, i.e. non-necessity ($\neg\Box p$) and impossibility ($\neg\Diamond p$). Our data suggest that in several typologically unrelated languages, negative modality generally involves separate, overt expression of negation and a modal in any force/flavor combination except deontic impossibility (\approx prohibition). We illustrate this pattern in Hausa, Thai and Kîtharaka, propose a (somewhat tentative) semantic analysis of deontic impossibility in these languages, and briefly outline how our observations may be captured in terms of optimization of the informativeness/complexity trade-off (see e.g. Kemp et al. 2018; Steinert-Threlkeld 2019; Uegaki 2023).

1 Introduction

In this paper, we present results from a crosslinguistic elicitation study on the expression of modal meaning. Broadly speaking, we discuss negative modality, i.e. non-necessity and impossibility, as illustrated in (1) and (2) with examples from English.

- (1) Non-necessity
 - a. I **don't have to** wear formal clothes to the party. (deontic)
 - b. You **need not** take the bus to get to the shopping center. (teleological)
- (2) Impossibility
 - a. You **mustn't** ride a motorbike without a helmet. (deontic)
 - b. You **can't** take another road to get to the city. (teleological)

From a wider perspective, we are interested in how modal meanings, including non-necessity and impossibility, are encoded across languages. To investigate this, we elicited data from the languages listed in (3) (grouped by language family where applicable).

- (3) **Language sample:** Dutch, Farsi, Greek, Hindi, Russian, Spanish (Indo-European), Vietnamese, Khmer (Austro-Asiatic), Thai (Kra-Dai), Cantonese (Sino-Tibetan), Hausa, Hebrew (Afro-Asiatic), Hungarian (Uralic), Akan, Igbo, Kîtharaka (Niger-Congo), Tagalog (Austronesian), Turkish (Turkic), Telugu (Dravidian), Japanese, Korean, Basque, Mapudungun

¹We would like to thank Ciyang Qing, Malte Zimmermann, as well as the reviewers and participants of TripleA for their feedback. Special thanks to our language consultants for providing and discussing the data presented in this paper. Ella Hannon has also contributed to an early stage of the data collection as a research assistant. All mistakes and omissions are ours.

Regarding negative modality in particular, our elicitations revealed that negative modality generally involves overt realization of negation and a modal, as in the English examples in (1) and (2). However, there is an interesting exception to this in several unrelated languages in our sample, where the meaning of deontic impossibility (\approx prohibition, see (2-a)) is lexicalized or otherwise less transparently encoded. In this paper, we will focus on three of these languages: Hausa (Chadic, Afro-Asiatic), Thai (Tai, Kra-Dai) and Kîtharaka (Bantu, Niger-Congo), and have a closer look at the relevant deontic impossibility constructions. The paper is structured as follows. Sect. 2 provides a brief overview of our elicitation methods and presents our central empirical observations. Sect. 3 constitutes the core of the paper. In Sect. 3.1, we briefly discuss the possible coding strategies for deontic impossibility that we are interested in, setting the stage for more detailed discussion of the deontic impossibility expressions that we found in Hausa, Thai and Kîtharaka. Based on previous works and our own data, we will propose a concrete analysis for deontic impossibility in Hausa in Sect. 3.2. Inspired by a similar-looking necessity construction as analyzed by Grubic and Mucha (2021), we start from the idea that deontic impossibility can be decomposed into a negative operator and a covert possibility modal. We ultimately reject this decompositional analysis for Hausa and propose that in Hausa (and in Thai) deontic impossibility is lexicalized in a single element. In Sect. 3.3, however, we show that such a decompositional account might be the right approach to analyzing deontic impossibility in Kîtharaka. Sect. 4 zooms out again and gives an outlook on how our overall crosslinguistic observations on the lexicalization of negative modality meanings can be explained. Sect. 5 concludes the paper.

2 Data and methodology

We elicited data from the languages in (3) with an adapted version of Vander Klok’s (2021) revised modal questionnaire for crosslinguistic use. This questionnaire is constructed for eliciting comparable data on the expression of various combinations of modal force and modal flavor (as defined in Kratzerian accounts of modality), and is therefore particularly suitable for detecting interesting crosslinguistic patterns. As will be illustrated in the course of this section, we adapted some of the materials in Vander Klok (2021) to be suitable to elicit negative modality, i.e. *non-necessity* ($\neg\Box p$) and *impossibility* ($\neg\Diamond p$) with epistemic, deontic, teleological, and pure circumstantial flavor.

The original data in this paper are based on elicitations with one native speaker per language. The consultants have a background in linguistics, and English was used as a language of communication throughout. The elicitations were conducted online in video conference calls, with each elicitation session jointly led by two of the authors. We used the modality questionnaire mainly for ‘translation in context’-tasks, i.e. speakers were presented with a context description and an English target sentence that contains a modal that is suitable in the context. The consultants’ task was to translate the target sentence into their native language such that it is equally felicitous in the provided context. Beyond these translation tasks, we also elicited acceptability judgments on selected examples and constructions. Our central empirical generalizations on the expression of negative modality are summarized in (4).

- (4) a. **Non-necessity** meaning is always realized as a combination of morphologically overt negation and a modal marker in all the languages in our sample.

- b. **Impossibility** meaning, by contrast, is lexicalized or otherwise less transparently encoded in the following languages in our sample: Basque, Hausa, Hebrew, Hungarian, Kîtharaka, Russian, Thai, and Turkish.

Within the subset of languages in (4-b), we observe an interesting division: Basque and Turkish have a lexicalized impossibility modal that is compatible with all modal flavors. In the other languages in (4-b), we found a specific lexical element or coding strategy for impossibility that is only compatible with deontic flavor. In what follows, we focus on three of these languages: Hausa, Thai and Kîtharaka. Let us illustrate the data pattern with some examples², beginning with non-necessity. Like all languages in the sample, Hausa, Thai and Kîtharaka express non-necessity with the combination of morphological negation and a clearly identifiable necessity modal. In Hausa (5-a), the necessity modal *dole* occurs with the parenthetic negation *ba...ba*. In Thai and Kîtharaka, the negation markers (*mai* and *ti*, respectively) precede the respective necessity modals. While we only illustrate this for the case of deontic non-necessity in (5), the generalization in (4-a) holds for non-necessity in all modal flavors. In all examples in this section, the English target sentences (boldfaced in the context descriptions) are provided as presented to the consultants, in lieu of translations below the example sentences.

- (5) *Context (deontic non-necessity)*: Your friend is having a birthday party. She tells you that she would like it if everyone invited wore formal clothes, but that it is not necessary to wear formal clothes to attend. You think to yourself ... I ought to wear formal clothes to the party, but **I don't have to wear formal clothes.**
- a. **Ba dole ba** ne in saka kayan bukin.
NEG MOD(□) NEG COP 1SG wear clothes party (Hausa)
- b. Chan **mai jambpen** daeng dour supbarb.
1SG NEG MOD(□) dress body formal (Thai)
- c. **Ti racima** wîkîre nguô cia ûbici.
NEG MOD(□) SM.wear.SBJV clothes official (Kîtharaka)

The most interesting case for our purposes is that of deontic impossibility. In Hausa (6-a) and Thai (6-b), this meaning seems to be encoded in a single lexical item, i.e. *kada* in Hausa and *harm* in Thai. In Kîtharaka (6-c), we see morphologically transparent negation (again encoded by the morpheme *ti*) combining with a subjunctive form of the verb, but there is no overt modal.

- (6) *Context (deontic impossibility)*: You are going to visit your friend in the hospital. When you enter into the hospital, you stop at the information desk to inquire what room your friend is in. But the woman at the information desk tells you that you can't visit your friend now because it's already 8pm. She says: "I'm sorry, the hospital regulations say that ... **Visitors mustn't stay after 6pm.**"
- a. **Kada** maziyata su wuce karfe 6 na yamma.
KADA visitors 3PL.SBJV stay hour 6 pm (Hausa)

²The following glosses are used in the examples: and = andative, f = feminine, foc = focus, fut = future, fv = final vowel, infv = infinitive, ipfv = imperfective, m = masculine, mod = modal, neg = negation, pfv = perfective, pl = plural, pres = present, pst = past, rel = relative form, sbjv = subjunctive, sg = singular, sm = subject marker.

- b. Yardpubuay **harm** yuu lang hok mornɔŋ.
patient.visitor HARM exist after six o'clock (Thai)
- c. Ageni ba-**ti-ka**-kinyithi-e thaa kûmi na ciîrî kûkû.
visitors SM-NEG-FUT/SBJV-reach-FV.SBJV hours ten and two in.here (Kîîtharaka)

As noted above, in Hausa, Thai and Kîîtharaka, the realizations in (6) are specific to deontic impossibility. Consider, by contrast, the examples in (7). To express the epistemic meaning specified in the context, all three languages use a combination of the same negation marker as in (5) with a possibility modal (*yiwu* in Hausa, *dai* in Thai, *ûmba* in Kîîtharaka). Thus, epistemic impossibility is encoded in a transparent manner, similar to the encoding of non-necessity. The impossibility constructions that we elicited in the deontic case in (6) are infelicitous in an epistemic impossibility context, as shown in (7-a-ii), (7-b-ii) and (7-c-ii).

(7) *Context: (epistemic impossibility):* Ben goes swimming every day. Ben is not obliged or required to go swimming; it is just a habit of his. It is now time for Ben to be swimming, so ... **Ben can't be at home.**

- a. (i) **Ba zai yiwu** Ben ya kasance a gida **ba**.
NEG 3SG.FUT MOD(◇) Ben 3SG be at house NEG
- (ii) **#Kada** Ben ya kasance a gida.
KADA Ben 3SG.M.SBJV be at house (Hausa)
- b. (i) Ben yuu tii baan **mai dai**.
Ben exist LOC house NEG MOD(◇)
- (ii) **#Ben harm** yuu tii baan.
Ben HARM exist LOC house (Thai)
- c. (i) Ben a-**ti-ûmba** kwîgua arî mûciî.
Ben SM-NEG-MOD(◇) INF.be SM.be home
- (ii) **#Ben a-ti-ga-kar-e** mûcii.
Ben SM-NEG-FUT/SBJV-stay-FV.SBJV home (Kîîtharaka)

With the context in (8), we illustrate the same point for teleological impossibility. The observations are identical to the epistemic impossibility case in (7) (except that Hausa uses different possibility modals to express the respective flavors).

(8) *Context (teleological impossibility):* There is only one main road from Location A to Location B ... If you are going from Location A to Location B, **you can't take another road.**

- a. (i) **Ba za ka iya** bin wata hanyar **ba**.
NEG FUT 2SG.M.SBJV MOD(◇) take another road NEG
- (ii) **#Kada** ka (iya) bin wata hanyar.
KADA 2SG.M.SBJV (MOD(◇)) take another road (Hausa)
- b. (i) Chay tanon anern **mai dai**.
use road another NEG MOD(◇)

- (ii) #Ben **harm** yuu tii baan.
Ben HARM exist LOC house (Thai)
- c. (i) \hat{U} -**ti-umba** gûtômîrîa njîra yîngî kîuthi B.
2SG-NEG-MOD(\diamond) use.PRES path other go B
- (ii) # \hat{U} -**ti-ga-tûmîr-e** njîra yîngî kîuthi B.
2SG-NEG-FUT/SBJV-use-FV.SBJV path other go B (Kîîtharaka)

Thus, it appears that there is something particular about deontic impossibility in these and other languages that leads to specific coding strategies for expressing this meaning. In this paper, we will only very briefly touch on the question of what might set deontic impossibility apart in that way (see Sect. 4). We also note, however, that semantic analyses of negative modality are in short supply, even more so when it comes to underrepresented languages. For this reason, we want to take a closer look at the deontic impossibility constructions in Hausa, Thai and Kîîtharaka, and discuss how they may be analyzed in a formal semantic framework.

3 A closer look at deontic impossibility

3.1 Strategies of expressing deontic impossibility

Before zooming in on the deontic impossibility constructions in the languages under investigation, let us briefly discuss how languages may encode prohibitive meanings, and define the range of crosslinguistic variation we expect. Firstly, there are two logical possibilities in which modal and negative meaning can combine to derive prohibition, sketched in (9-a) and (9-b). In (9-a), a negative operator scopes over a possibility modal. Talking about deontic impossibility in particular, this constellation would roughly correspond to *not allowed* in English. However, the same semantic effect is achieved if a necessity modal scopes over negation, as sketched in (9-b). This is exemplified by *mustn't* in English.

- (9) a. [NEG [\diamond [p]]] (\approx 'not allowed')
b. [\square [NEG [p]]] (\approx 'mustn't')

As in English both the negative and the modal meaning components are realized overtly, we would speak of a transparent coding strategy in this case. An example of a non-transparent coding strategy would be a construction in which only negation is overtly realized, but the overall sentence has a modal meaning. If a covert modal operator is assumed in the analysis of such a construction, the question arises whether modality is contributed by a possibility modal as in (9-a) or a necessity modal as in (9-b). Interestingly, this question has been discussed in the literature with reference to English examples such as (10).

- (10) No walking on the grass! (Iatridou, 2021, 520)

Iatridou (2021) refers to constructions of this kind as 'negation-licensed commands', and proposes an analysis along the lines of (9-b) with a covert necessity modal. However, the same type of construction has received analyses in the spirit of (9-a), featuring a covert or elided deontic possi-

bility modal scoping under negation (see Donovan 2020, Frühauf et al. 2023). Finally, the negative modal meanings in (9) can also be bundled in a single lexical element, as schematized in (11).

- (11) a. $[\neg \diamond [p]]$
 b. $[\square \neg [p]]$

Arguably, it would be very difficult to distinguish (11-a) and (11-b) empirically. For the purposes of this paper, we will not discuss the distinction between (11-a) and (11-b) any further, but subsume both under the cover of lexicalized impossibility and assume the semantics in (11-a). In the remainder of this section, we discuss how the non-transparent strategies of encoding prohibition that we introduced above apply to Hausa, Thai and Kîtharaka.

3.2 Deontic impossibility in Hausa

On the way towards an analysis of deontic impossibility in the languages under consideration, we take Hausa as a starting point. In particular, we consider the necessity construction in (12), which is discussed in detail in Grubic and Mucha (2021).

- (12) *Deontic necessity*: According to the Nigerian law, ...
 ... **sai** Audu **ya** tafi fuřsuna.
 SAI Audu 3SG.M.SBJV go prison
 ‘Audu must go to jail.’

In (12), the exclusive particle *sai* combines with the TAM form that is labeled as ‘subjunctive’ in descriptive works on Hausa. The interpretation of (12) is deontic necessity. Crucially, the modal use of *sai* only occurs with this particular TAM form. With other TAMs, *sai* behaves as a standard exhaustive operator akin to English *only*, excluding alternatives to the focused constituent it associates with. The examples in (13) and (14) (taken from Grubic and Mucha 2021, 361) illustrate this for direct object focus and sentence focus, respectively.

- (13) Sai **tuwō** mātā sukè girkà.
 SAI fufu women 3PL.REL.IPFV cook
 ‘The women are only cooking FUFU.’ (object focus)
- (14) *Sentence focus*: Did something happen?
 Ā’à, sai **mātā sun girkà tuwō**.
 no SAI women 3PL.PFV cook fufu
 ‘No, except that the women cooked fufu.’ (sentence focus)

Grubic and Mucha (2021) propose a decompositional analysis of the construction in (12), with the following ingredients: i) *sai* retains its function as an exclusive operator, and it associates with the whole sentence, ii) the so-called ‘subjunctive’ TAM in Hausa is actually a prospective aspect, i.e. it denotes future-shifting of event times³, iii) modality in sentences like (12) is contributed by a covert possibility operator. The sentence in (12) thus has the structure in (15-a). *Sai* associates

³This idea was formalized and defended in Mucha (2013), and inspired by an earlier discussion in Schuh (2003).

with future possibilities, and ends up excluding all such possibilities that are not entailed by the prejacent. The truth conditions that this analysis derives for (12) can be paraphrased as (15-b).

- (15) a. [SAI [\diamond [PROSP [Audu go to jail]]]]
 b. \approx The only future possibility is for Audu to go to jail.

This analysis is crucially inspired by Kaufmann (2012), who proposes that the necessity interpretation of imperatives should be accounted for in terms of exhaustification of possibilities. While *sai* overtly contributes exhaustification in cases like (12), in imperatives exhaustification is covert. As will become relevant later in our discussion of deontic impossibility in Kîtharaka (Sect. 3.3), Grubic and Mucha (2021) present data suggesting that exhaustification can be covert in the relevant necessity constructions in Hausa as well. They present the minimal pair in (16-a) and (16-b), differing only in whether or not *sai* is present in the structure. They note that both sentences obtain the same modal necessity interpretation, and propose a covert EXH operator for (16-b).

- (16) a. Sai sù tunà sañai.
 SAI 3PL.SBJV remember well
 [SAI [\diamond [PROSP [they remember]]]]]
 b. Sù tunà sañai.
 3PL.SBJV remember well
 [EXH [\diamond [PROSP [they remember]]]]]
 ‘They must/should remember.’ (adapted from Grubic and Mucha 2021, 366)

The reason we are taking examples such as (12) and (16-a) as a starting point for our discussion of deontic impossibility is that, at least in Hausa, the respective constructions look remarkably similar in terms of their surface structure. Recall the deontic impossibility construction in (6-a), repeated below with the context abridged. Similar to necessity constructions with *sai*, the modal sentence combines a sentence-initial particle (i.e. *kada* in (6-a)) with ‘subjunctive’ TAM marking.

- (6-a) *Deontic impossibility*: ... “I’m sorry, the hospital regulations say that ...”
kada maziyarƙa **su** wuce ƙarfe 6 na yamma.
 KADA visitors 3PL.SBJV stay hour 6 pm
 ‘... visitors mustn’t stay after 6pm.’ (Hausa)

The structural similarity of the two constructions suggests that deontic impossibility in Hausa could receive a decompositional analysis in parallel to the account of *sai* constructions proposed by Grubic and Mucha (2021). Under such an account, *kada* would combine with prospective possibilities (just like *sai* does), but would negate, rather than exhaustify, these possibilities. The logical structure of (6-a) would then look like (17-a), and the meaning of the sentence could be paraphrased as (17-b).

- (17) a. [NEG [\diamond [PROSP [visitors stay 6pm]]]]
 b. \approx It is not possible that visitors stay after 6pm.

In the remainder of this subsection, we further explore this potential analysis of the deontic impossibility construction. Previewing the results of this discussion, we will ultimately discard such a

with all non-deontic modal flavors.⁴ The incompatibility of the *kada*-construction with epistemic and teleological flavors has already been illustrated in Sect. 2 (see examples (7-a-ii) and (8-a-ii)). In (20), we show that *kada* cannot express circumstantial impossibility, either.

(20) *Context (circumstantial impossibility)*: Ben was in a motorbike accident 3 weeks ago. He sprained his ankle and is in a lot of pain. No one is forbidding him to walk, but he is not able to, because his ankle hurts so much.

#**Kada** Ben ya tafi.
 KADA Ben 3SG.M.SBJV walk
 Intended: ‘Ben can’t walk.’

From these data, we tentatively conclude that the modal constructions with *sai* and *kada* differ in their restrictions on modal flavor. This can be viewed as additional evidence showing that the two constructions are not entirely parallel, inasmuch as they could not involve the exact same possibility operator.

To sum up the discussion thus far, we considered a fully decompositional analysis of deontic impossibility in Hausa, based on what Grubic and Mucha (2021) propose for a seemingly similar necessity construction involving the exclusive particle *sai*. We then sketched three arguments against such a parallel analysis: i) the relevant sentence-initial element *kada* is described as a prohibitive, rather than a purely negative marker, in previous literature (Newman, 1971, 2000), ii) we have found no independent evidence for non-modal negative uses of *kada*, iii) the *kada*-construction and the *sai*-construction seem to differ in their restrictions on modal flavor. Hence, rather than decomposing deontic impossibility into negation (encoded by *kada*) and covert possibility, we propose that *kada* lexicalizes both meaning components. Thus revising the working hypothesis in (17) for Hausa, we propose that the deontic impossibility statement in (6-a) has the structure in (21), with *kada* contributing negative modality.

(6-a) *Deontic impossibility*: ... “I’m sorry, the hospital regulations say that...”

kada maziyar^{ta} **su** wuce ^{ƙarfe} 6 na yamma.
 KADA visitors 3PL.SBJV stay hour 6 pm
 ‘visitors mustn’t stay after 6pm.’ (Hausa)

(21) [¬◇ [PROSP [visitors stay 6pm]]]

In (22) we provide a concrete lexical entry for *kada* as a deontic impossibility modal, much in the spirit of Kratzer (1981) and countless subsequent works on natural language modality. According to (22), the meaning of *kada* differs from other deontic modals only in that its asserted meaning component contains negation.

(22) $[[kada]]^c$ is only defined if *c* provides a **deontic ordering source** *O* and a circumstantial modal base *MB*, if defined: $[[kada]]^c = \lambda p. \lambda w. \neg \exists w' [w' \in \text{BEST}_{O(w)}(\text{MB}(w)) \ \& \ p(w')]$

⁴While the judgments seem quite clear for the main clause uses of *kada* that we tested with our consultant, it is conceivable that the ‘lest’-construction illustrated in (18) should be viewed as an instance of teleological impossibility. We will leave this issue unresolved here.

Based on our own data and the previous literature on Kĩĩtharaka, the status of *ti* as a negation marker seems uncontroversial. The crucial challenge thus amounts to correctly analyzing the meaning contribution of *ka* in constructions like (6-c), and to understand the construction as a whole. While we cannot solve this issue conclusively in the present paper, we will progress by relating our original data to some relevant discussion in the previous literature.

First, let us zoom out and consider some pertinent generalizations on languages of the Bantu family. Interestingly, our observation that in Kĩĩtharaka prohibition is expressed by the combination of negation and a subjunctive verb form appears to be somewhat representative of Bantu languages. Devos and Van Olmen (2013) investigated the grammatical marking of imperative and prohibitive meanings in a sample of over 100 Bantu languages. The majority of these languages use the combination of a negation marker and subjunctive morphology to express prohibition.⁶ Even the use of a pre-verbal marker *ka* in prohibitive constructions is attested in other Bantu languages. Devos and Van Olmen (2013) mention Kamba and Gĩkũyũ as further examples of the same kind of constructions we elicited in Kĩĩtharaka, and provide the example in (24) (taken from Barlow 1951) for illustration. Notably, they categorize this *ka*-marker as an adative/motional prefix and do not mention any modal (or temporal) meaning that this morpheme might contribute.

- (24) mu-**ti-ka**-gwat-e
 SM-NEG-AND-take.hold-SBJV
 ‘Don’t take hold!’ (Gĩkũyũ, Devos and Van Olmen 2013, 27)

Even more pertinently, Muriungi (2005) provides some discussion of the *ka*-marker in Kĩĩtharaka in particular. Muriungi explicitly classifies *ka* as a future marker, giving examples such as (25).

- (25) Karimi a-**ka-rug-a** katoroko
 Karimi SM-FUT-cook-FV beer
 ‘Karimi will prepare Katoroko.’ (Kĩĩtharaka, adapted from Muriungi 2005, 47)

Beyond this, Muriungi (2005) presents a number of further generalizations that are potentially relevant for our purposes. For instance, Muriungi states that when a focus marker is prefixed to the verbal complex in (25), the sentence obtains a modal necessity reading, as shown in (26). Notably, neither (25) nor (26) have the subjunctive final vowel. Hence, howsoever modal readings are derived in the relevant Kĩĩtharaka examples seems to be at least partially independent of mood.

- (26) Karimi **n-a-ka-rug-a** katoroko
 Karimi FOC-SM-FUT-cook-FV beer
 ‘Karimi must prepare Katoroko.’
 NOT: ‘Karimi will prepare Katoroko.’ (Kĩĩtharaka, adapted from Muriungi 2005, 47)

Muriungi’s (2005) notes on the negative marker *ti* are interesting as well. According to the author, when future sentences such as (25) are negated, the *ka*-marker is omitted:⁷

⁶It should be noted that Devos and Van Olmen (2013) talk about prohibition in a narrower sense than we do, focusing on negative imperatives with second person subjects.

⁷This generalization might seem puzzling, and we cannot provide an analysis for the case of Kĩĩtharaka. Note however that negation, and more broadly non-veridical operators, have been claimed to license future interpretations in the absence of overt future marking in some languages (see e.g. Mucha (2016) on the Grassfields Bantu language

- (27) Karimi a-**ti**-(***ka**)-rug-a kathoroko
 Karimi SM-NEG-FUT-cook-FV beer
 ‘Karimi will not prepare Kathoroko.’ (Kîtharaka, adapted from Muriungi 2005, 54)

Muriungi also notes that, when *ti* does occur with *ka*, “it gives rise to a meaning of roughly the form ‘don’t’ ” (Muriungi, 2005, 54). The author illustrates this with the example in (28), which corresponds to the structures we elicited in deontic impossibility contexts in Kîtharaka.

- (28) U-**ti-ka**-rongo-e
 2SG-NEG-FUT/SBJV-cheat-SBJV
 ‘Don’t cheat!’ (Kîtharaka, adapted from Muriungi 2005, 54)

Thus, a relevant generalization seems to be that *ka* obtains clearly modal readings when combining with a focus marker or with negation. In this context, it is particularly interesting to note that, again according to Muriungi (2005), these two morphemes are in complementary distribution:

- (29) *Paul n-a-**ti/ta**-ra-rug-a nkima
 Paul FOC-SM-NEG-PST-cook-FV food
 Intended: ‘Paul did not cook food.’ (Kîtharaka, adapted from Muriungi 2005, 80)

We submit that Muriungi’s generalizations provide some support for the idea that deontic impossibility can be decomposed in Kîtharaka.

Let us first look at the necessity sentence in (26), to see whether it might lend itself to a decompositional analysis along the lines of Grubic and Mucha’s (2021) analysis of the Hausa *sai*-construction. To make the analogy with the Hausa data explicit: we conjecture that Kîtharaka necessity constructions such as (26) have the same logical structure as Hausa sentences such as (16-b), in which the exclusive particle *sai* is replaced by a covert exhaustivity operator. (The example is repeated below for convenience.)

- (16) a. Sai sù tunà sañai.
 SAI 3PL.SBJV remember well
 [SAI [◇ [PROSP [they remember]]]]
 b. Sù tunà sañai.
 3PL.SBJV remember well
 [EXH [◇ [PROSP [they remember]]]]
 ‘They must/should remember.’
 (Hausa, adapted from Grubic and Mucha 2021, 366)

Of course, the structure proposed for (16-b) has some ingredients that we do not have direct evidence for on the surface, namely a covert exhaustivity operator that associates with the whole sentence and scopes over a likewise covert possibility operator. Let us see how plausible it is to stipulate these in the case of Kîtharaka. Firstly, Muriungi (2005) shows that pre-verbal focus marking indicates focus on the VP or the whole sentence, as opposed to subject or object focus (in which case the focus marker is prefixed to the respective nominal constituent). Against this

Medumba, and Bochnak (2019) for a more general discussion of the phenomenon).

In our analytical sketch above, we have worked on the assumption that the morpheme *ka* encodes prospectivity, in parallel to the so-called ‘subjunctive’ TAM form in Hausa. However, it is well-known that future markers in many languages are associated with modality as well, and in this case are usually analyzed as universal quantifiers over possible worlds (see Enç 1996; Copley 2009; Tonhauser 2011; Mucha 2016; Giannakidou and Mari 2018, among many others). If *ka* denoted a modal future of this sort, it would contribute a type of necessity meaning, and, if *ka* could be argued to outscope the negation marker at LF, the Kĩtharaka impossibility construction could be analyzed as a ‘negation-licensed command’ in the sense of Iatridou (2021). The choice between these possible analyses depends on further, more detailed investigation of the ‘future marker’ *ka* in Kĩtharaka.

4 Outlook

In Sect. 3, we argued that deontic impossibility is lexicalized in Hausa and Thai. Moreover, recall from Sect. 2 that we found lexicalized impossibility in several more languages in our language sample, and that the lexicalized impossibility modals in these languages were either compatible with all modal flavors, or with deontic flavor only. These findings from our overall study can be summarized as in Table 1 below (we do not include Kĩtharaka here, since deontic impossibility is not lexicalized in this language).

	Non-necessity (any flavor)	Impossibility		
		epistemic	deontic	other root flavors
Basque, Turkish	×	✓	✓	✓
Hausa, Hebrew, Thai Hungarian, Russian	×	×	✓	×

Table 1: ✓ means the meaning is lexicalized, × means it is not

One way to look at these results is that all the lexicalized impossibility modals that we encountered in our crosslinguistic study are compatible with deontic flavor, and they differ in whether they can express other modal flavors in addition. Based on this observation, Uegaki et al. (in progress) hypothesize the generalization formulated in (30).

- (30) **Deontic Priority (DP) generalization:** If a language lexicalizes any impossibilities, then it lexicalizes deontic impossibility.

While we have discussed the observed asymmetry between deontic possibility and other flavors of impossibility in some detail for the cases of Hausa, Thai and Kĩtharaka, our discussion in the previous section does not provide an explanation for this asymmetry. Although a detailed explanation is beyond the scope of this paper and will be left for another occasion (Uegaki et al., in progress), we will at least outline in what direction such an explanation may go.

We propose that Deontic Priority might be rooted in grammar-external communicative pressures. A number of relevant recent proposals have captured lexicalization patterns in logical vocabularies, broadly in terms of informativeness and/or complexity (e.g., Kemp et al. 2018; Imel and

Steinert-Threlkeld 2022; Uegaki 2023). The DP generalization suggests that a theory of modal lexicalization must capture the contrast between flavors (deontic vs. others) in some way. In a computational modeling study, Uegaki et al. (in progress) explore how Imel & Steinert-Threlkeld's (2022) existing model in terms of a complexity/informativeness trade-off can be extended to capture the contrast. In particular, the study explores if asymmetries in the communicative utility function yield a picture in which a utility bias in favor of deontic impossibility correlates with the optimality of languages that satisfy the DP generalization. The results suggest that optimizing the trade-off between simplicity and informativeness, in the presence of a bias for the deontic flavor, could indeed explain the DP generalization. The intuition that this model is intended to capture is fairly simple, and has been expressed in some form in the typological literature before (e.g. Aikhenvald 2010, see also Devos and Van Olmen 2013): in the case of deontic impossibility, speakers face a particularly high pressure to communicate the intended meaning successfully, since the communicative function of prohibition is to prevent negative and potentially dangerous situations. If the hypothesis in (30) (or some version of it) is correct, this functional prominence of deontic impossibility is reflected in the modal systems of some languages as more efficient and morphologically simpler surface realization when compared to other negative modal meanings.

5 Conclusions

In this paper, we discussed negative modality from a crosslinguistic perspective, based on original data elicited from 23 languages. We started from the general observation that negative modal meanings show variation in how they are encoded in our language sample. While non-necessity always involves overt realization of negation and a modal operator,⁹ impossibility is lexicalized or otherwise less transparently encoded in some languages. We had a closer look at impossibility in three languages in which these particular forms are restricted to deontic impossibility meaning (i.e. prohibition), and discussed how the respective constructions can be analyzed.

From a broader perspective, the results of our crosslinguistic study suggest an asymmetry between prohibition and other flavors of impossibility—we did not find any impossibility modal that is not compatible with deontic flavor. In a brief outlook, we suggested that this finding might point to a potentially universal crosslinguistic generalization, the ‘Deontic Priority generalization’. We also hinted at the idea that this generalization can be captured under the assumption that optimization of the simplicity/informativeness trade-off in natural languages might involve a utility bias for deontic impossibility meaning, an idea that is being fleshed out in ongoing research (Uegaki et al., in progress).

⁹While we did not discuss this finding in the present paper, we submit that an explanation in the same spirit as laid out in Sect. 4 can account for it as well. In particular, assuming that non-necessity represents the ‘O-corner’ in the Aristotelian square of oppositions (see Horn 1989), the lack of lexicalized non-necessity is to be expected.

References

- Abels, K. and P. Muriungi (2008). The focus marker in Kĩtharaka: Syntax and semantics. *Lingua* 118(5), 687–731.
- Aikhenvald, A. (2010). *Imperatives and commands*. Oxford: Oxford University Press.
- Barlow, A. R. (1951). *Studies in Kikuyu grammar and idiom*. Edinburgh: William Blackwood & Sons.
- Bochnak, M. R. (2019). Future reference with and without future marking. *Language and Linguistics Compass* 13(1), e12307.
- Copley, B. (2009). *The semantics of the future*. New York, London: Routledge.
- Devos, M. and D. Van Olmen (2013). Describing and explaining the variation of Bantu imperatives and prohibitives. *Studies in Language* 37(1), 1–57.
- Donovan, M. (2020). General prohibition. Ms., University of Delaware.
- Enç, M. (1996). Tense and Modality. In S. Lappin (Ed.), *The Handbook of Contemporary Semantic Theory*. London: Blackwell.
- Frühauf, F., H. Karawani, T. Koev, N. Korotkova, D. Penka, and D. Skibra (2023). Vanilla rules: The “no ice cream” construction. In M. Onoeva, A. Staňková, and R. Šimík (Eds.), *Proceedings of Sinn und Bedeutung* 27, pp. 209–227.
- Giannakidou, A. and A. Mari (2018). A unified analysis of the future as epistemic modality: The view from Greek and Italian. *Natural Language & Linguistic Theory* 36(1), 85–129.
- Grubic, M. and A. Mucha (2021). Decomposing Necessity – The Hausa exclusive particle *sai* as a window into the building blocks of modal meaning. In P. Grosz, L. Martí, H. Pearson, Y. Sudo, and S. Zobel (Eds.), *Proceedings of Sinn und Bedeutung* 25, pp. 358–375.
- Grubic, M., A. Renans, and R. A. Duah (2019). Focus, exhaustivity and existence in Akan, Ga and Ngamo. *Linguistics* 57(1), 221–268.
- Horn, L. R. (1989). *A natural history of negation*. Chicago: University of Chicago Press.
- Iatridou, S. (2021). Negation-licensed commands. *Linguistic Inquiry* 52(3), 519–549.
- Imel, N. and S. Steinert-Threlkeld (2022). Modal semantic universals optimize the simplicity/informativeness trade-off. In *Proceedings of Semantics and Linguistic Theory (SALT)* 32, pp. 227–248.
- Kaufmann, M. (2012). *Interpreting Imperatives*, Volume 88 of *Studies in Linguistics and Philosophy*. New York: Springer.
- Kemp, C., Y. Xu, and T. Regier (2018). Semantic typology and efficient communication. *Annual Review of Linguistics* 4(1), 109–128.
- Kratzer, A. (1981). The notional category of modality. In H. J. Eikmeyer and H. Rieser (Eds.), *Words, Worlds and Contexts*. Berlin, New York: De Gruyter.
- Mucha, A. (2013). Temporal interpretation in Hausa. *Linguistics and Philosophy* 36(5), 371–415.
- Mucha, A. (2016). Deriving the temporal properties of future markers from aspect. In N. Bade, P. Berezovskaya, and A. Schöller (Eds.), *Proceedings of Sinn und Bedeutung* 20, pp. 533–550.
- Muriungi, P. (2005). Wh-questions in Kitharaka. *Studies in African Linguistics* 34(1), 43–104.
- Newman, P. (1971). The Hausa negative markers. *Studies in African Linguistics* 2(3), 183–195.
- Newman, P. (2000). *The Hausa language: An encyclopedic reference grammar*. New Haven: Yale University Press.

- Ramchand, G. C. (2018). *Situations and syntactic structures*, Volume 77 of *Linguistic Inquiry Monographs*. Cambridge, MA, London: The MIT Press.
- Schuh, R. G. (2003). The functional unity of the Hausa and West Chadic subjunctive. In *UCLA Working Papers in Linguistics* 9, pp. 17–42.
- Steinert-Threlkeld, S. (2019). Quantifiers in natural language optimize the simplicity/ informativeness trade-off. In J. J. Schläder, D. McHugh, and F. Roelofsen (Eds.), *Proceedings of the 22nd Amsterdam colloquium*, pp. 513–522. University of Amsterdam.
- Tonhauser, J. (2011). The Paraguayan Guaraní future marker *-ta*: formal semantics and cross-linguistic comparison. In M. Rathert and R. Musan (Eds.), *Tense Across Languages*, pp. 207–231. Tübingen: Niemeyer.
- Uegaki, W. (2023). The informativeness/complexity trade-off in the domain of Boolean connectives. *Linguistic Inquiry* 55(1), 174–196.
- Uegaki, W., A. Mucha, N. Imel, and S. Steinert-Threlkeld (in progress). Deontic priority in the lexicalization of impossibility modals. Ms., University of Edinburgh, UC Irvine, University of Washington, preliminary version available at <https://psyarxiv.com/h63y9/>.
- Vander Klok, J. (2021). Revised modal questionnaire for cross-linguistic use. https://www.eva.mpg.de/~lingua/tools-at-lingboard/pdf/~Modal/~_Questionnaire/~_CrossLing/~JVK.pdf.

Two Types of Definiteness: A Case Study from Hakha Lai¹

James C. Wamsley – Indiana University

Abstract. The question of how to formally represent definites as a class of referential expressions has been a longstanding area of inquiry in linguistic research (Hawkins 1978; Heim 1982). This talk contributes to this research by analyzing elicitation data on definite expressions from four diaspora (Indianapolis) speakers of Hakha Lai, a South Central (formerly Kuki-Chin) Tibeto-Burman language (Chin State, Burma). While descriptive accounts of definiteness marking in Lai exist (Barnes 1998; Baclawski 2013), this talk contributes novel data as well as a fine-grained syntactic analysis of referential expressions in Hakha Lai and contributes more generally to ongoing research on the role of indices and their status as syntactic objects.

1. Introduction

One of the ongoing discussions in research on nominals is how definite expressions are to be represented in formal semantics. This paper investigates the distinction between “unique” and “indexed” definites (Jenks & Konate 2022, Schwarz 2009). The analysis described here consults original data from four speakers of Hakha Lai, a South Central Tibeto-Burman (formerly known as Kuki-Chin) language, spoken in Chin state in Burma and by a large speaker community in southern Indianapolis, numbering in the tens of thousands (Salaz & Raymer 2020).² The analysis of Hakha Lai definite expressions provides support for the delineation between “unique” and “indexed” definites, appealing primarily to their structural differences. The basis of this analysis is the findings of recent studies of Washo (Hanink 2021), which found that indices, symbols which anchor a noun phrase to a particular discourse referent, are represented morphosyntactically within the expressions themselves. Similar behavior in Hakha Lai definite expressions is described here and contributes to ongoing investigations of the compositional role of indices in expressing definiteness in natural language.

Notably, in Hakha Lai, definite expressions often surface as bare nouns, meaning that there is no dedicated morpheme which can be categorized as a “definite article”. However, where overt morphological marking for definiteness is present, the delineation between “unique” and “indexed” definites is represented in the morphology by a morpheme *mah*, which occurs prenominal with demonstratives and as a suffix in personal pronouns. This paper is an investigation of the structural properties of definite expressions in Hakha Lai which contain *mah*.

¹ The author would like to thank the speakers who were generous with their time as well as the organizing committee of TripleA. Additional thanks go to Emily Hanink and Mary Moroney for their comments on the drafts.

² The data come from four speakers of Hakha Lai, each from different towns and cities within Chin State. The data sources are identified by the speakers’ hometowns, which are Hakha, Vawngtu, Thantlang, and Sunthla. Only data from the Hakha and Sunthla speakers appear in this article. Any data without a speaker reference comes from the researcher’s elicitation notes.

The paper continues as follows. Section 2 describes two types of definite expressions described in previous research, “unique” and “indexed” (terms from Jenks & Konate 2022). Section 3 describes definite expressions in Hakha Lai, showing that while definite expressions in Hakha Lai tend to be bare nouns, “indexed” expressions (including demonstratives and personal pronouns) contain a morpheme *mah*. Section 4 offers an analysis of Hakha Lai definite expressions, proposing a syntactic and semantic description of definites in general and for the *mah* morpheme in particular. Section 5 is a discussion of the analysis in Section 4 and its contribution to ongoing investigations of definiteness and the syntactic instantiation of indices (following Hanink 2021). Section 6 concludes.

2. Two categories of definite expressions

Recent typological investigations of definite expressions in various languages have yielded analyses which support claims that there are multiple kinds of definite expressions. One such delineation is the category of “unique” definites, a category of definites whose primary characteristic is the embedded presupposition of uniqueness for the given expression. An example of a “unique” definite is shown in (1) below.

(1) *Unique definite*

CONTEXT: There is one computer in the room.

[The computer] is new.

Analyses of “unique” definites claim that the uniqueness quality of the referent is what licenses the presence of a definite article. As we see in the context in (1), there is a single computer in the room that is salient to the speaker and addressee. Therefore, the expression *the computer* is appropriate for describing the unique referent. If there were more than one computer (or if there were no computer present), the expression *the computer* would fail to be interpretable by the addressee in the context.

The second relevant category for this investigation is the category of “indexed” definites. Recent studies of the unique/indexed definite split state that while “indexed” definite expressions still contain a presupposition of uniqueness, they differ from “unique” definites in that they also contain a semantic component which restricts their interpretation to a previously established discourse referent. This restriction is represented by an index, an indicator that the “indexed” definite expression is anchored to a prior discourse referent. An “indexed” definite expression is shown in example (2).

(2) *Indexed definite*

CONTEXT: Two friends are talking. One of them bought a computer and a phone. “Yesterday, I bought a new computer_i and a new phone. The phone didn’t cost too much but [the computer_i] was pretty expensive.”

3. Hakha Lai definite expressions

Hakha Lai (also referred to as ‘Hakha Chin’, ‘Lai’, ‘Lai Chin’, and ‘Laiholh’) is a South Central Tibeto-Burman language spoken in northwest Burma and in the Chin diaspora communities of the U.S. Hakha Lai is one of many Chin languages, ‘Chin’ being the self-designation of the ethnic group which speaks the South Central (formerly known as ‘Kuki-Chin’) languages spoken in Burma, Bangladesh, and India. Hakha Lai is the local variant spoken in Hakha, the capital of Chin State, and is a lingua franca in the Chin community, not only in Chin state, but also in the various diaspora communities worldwide and is thus the most widely spoken variant (Berkson et al. 2023).

Hakha Lai, like many other Tibeto-Burman languages, lacks a dedicated definite article.⁴ This means that definite expressions which in English would be accompanied by *the* would not have any additional morphology to distinguish it from an indefinite. This allows for ambiguous readings of sentences with bare nouns, as shown in the example below.

- (4) [uico cu] a-lian
 dog CU 3.SG-be.big
 ‘[The dog] is big.’ but also: ‘Dogs are big.’ (pilot data)

The next section discusses the category of unique definites in Hakha Lai. As a reminder, unique definites are those whose form is licensed by a presupposition of uniqueness alone. This category of definites differs from indexed definites in that they do not contain an index, anchoring them to a previously mentioned discourse referent.

3.1. Unique definites in Hakha Lai

Bare nouns can represent several kinds of sub-categories of definite expressions in Hakha Lai. For instance, example (5) below shows a bare noun used in an *immediate situation* context.

- (5) *Immediate situation definite*
CONTEXT: There is a baby sleeping. The speaker wants the addressee to stay quiet.
 Zangfang nak dai tein um uh. [Nau] aa-hngilh lio.
 please quiet ADV stay IMP baby 3.SG.REFL-sleep PROG
 ‘Please be quiet. [The baby] is sleeping.’ (Hakha speaker)

In example (5), the unique definite *nau* ‘the baby’ is a bare noun and is interpreted as referring to the unique baby which is present in the speaking context. Example (6) below provides an example of a bare noun in Lai used to represent a definite expression licensed by bridging. Bridging is a

⁴ However, see the discussion of *kha* in Section 4.

phenomenon wherein a definite expression is licensed even though the referent is not otherwise “present”, either in the discourse or the speaking situation.

(6) *Part-whole bridging definite*

CONTEXT: Jim and Dawn are talking about their friend Michelle. Michelle bought a new car, and the wheels are red. Jim comments on the car:

Michelle=nih	mawtaw	a-cawk.	[tire	cu]	a-sen
Michelle=ERG	car	3.SG-buy	tire	CU	3.SG-be.red
'Michelle bought a car. [The tires] were red.'					(Sunthla speaker)

In the example above, the expression *tire cu* is acceptable to refer to the tires which are part of the aforementioned car in the previous sentence.

Each of these examples where bare nouns are licensed are “unique” definites and do not identify a referent which is coreferential with another previously introduced discourse referent. The next several examples provide contexts in which “indexed” definites are often used.

3.2. Indexed definites in Hakha Lai

The following examples contain contexts in which indexed definites are expected to occur. Remember that while in English both unique and indexed definites surface with the same morphological marker (the definite article *the*), in other languages, there is either distinct marking for the two categories of definite expressions, or a distinct marker for one form (often indexed definites) while the other surfaces as a bare nominal.

3.2.1. Demonstratives

One category of definite expressions which are unambiguously indexed are demonstratives. In Hakha Lai, demonstrative expressions contain a pronominal element, *mah*, and a postnominal element, a “discourse deictic”. Discourse deictics belong to a category of postnominal markers which in demonstrative contexts encode spatial deixis. In non-demonstrative contexts, discourse deictics encode other discourse-level properties of nominals, such as temporal properties. The sentences in (7a-c) display the paradigmatic encoding of spatial deixis according to the postnominal discourse deictic.

(7) *Demonstratives in Hakha Lai*

a.	[mah	cauk	hi]	ka-rel	dih	ka te si
	DEM	book	SPKR.PROX	1.SG-read	COMP	just recently
	'I just finished reading this book (near me).'					(Hakha speaker)

expression contexts, demonstratives and anaphoric expressions, we already have seen that there are noticeable morphological differences from the unique definite expressions shown in Section 3.1. Let's next look at one more category of indexed definites, personal pronouns.

3.2.3. Personal pronouns

In Hakha Lai, there is a double personal pronoun system. The first set of pronouns contains person and number marking and ends with a suffix, *-mah*, which is comparable in form to the pronominal element seen in demonstratives. Hakha Lai is a pro-drop language, so it is rare that pronouns appear, but when they are used, it is often in the form seen in this first set of pronouns. The second set of pronouns are used in certain pragmatic situations, mostly emphatic or contrastive situations (Hlun 2007). The distinction between these forms requires further investigation. The two sets of pronouns are shown in Table 1. below, where Hlun uses the term “focus” pronouns to describe the common forms and “contrast” pronouns to describe the secondary forms.

First Person		
	Sg	Pl
Focus	keimah	kanmah
Contrast	kei	kannah

Second Person		
	Sg	Pl
Focus	nangmah	nanmah
Contrast	nang	nannah

Third Person		
	Sg	Pl
Focus	amah	anmah
Contrast	anih	annah

Table 1. Focus and Contrast pronouns (from Hlun 2007)

As shown in Table 1., each of the six pronouns encoding first, second, and third person singular and plural contains the suffix *-mah*. An example sentence with the pronoun *amah* ‘he’ is shown below.

- (9) *Hakha Lai personal pronoun*
 Café=ah saya=pa a-raa. [(Amah=pa)] luchin rang
 café=LOC teacher=MASC 3.SG-come 3.SG.PRO=MASC hat white
 a-chinh.
 3.SG-wear
 ‘A (male) teacher came into the café. [He] was wearing a white hat.’
 (Sunthla speaker)

In the example above, the personal pronoun is optionally present. Notably, like the other examples of indexed definites shown in sections 3.2.1 and 3.2.2, the morpheme *mah* is present in the morphological structure of the expression.

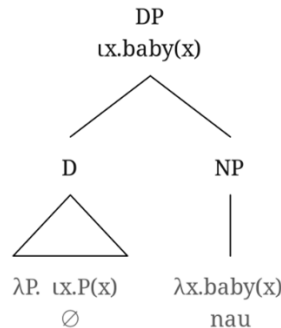
4. Analysis of Hakha Lai definite expressions

This section provides a semantic analysis of definite expressions in Hakha Lai. We will first discuss the semantics of unique definite expressions in Hakha Lai, which, like in other article-less languages, obtain their definite interpretation via a type-shifting iota operator ι (see Dayal 1999). Next, we will look at the category of indexed definite expressions and provide an analysis of the role of the morpheme *mah*, claiming that it is a syntactic instantiation of the index. This analysis will apply to all three types of definite expressions described above: demonstratives, anaphoric definites, and personal pronouns. This section is followed by a discussion of the analysis and how it compares with and contributes to ongoing investigations of definiteness in natural language.

4.1. The semantics of unique definites

Unique definites in Hakha Lai consist of a D head and an NP.⁵ The structure of the definite expression *nau* ‘the baby’ from example (5) above, is shown in (10) below.

(10) *Syntactic and semantic structure of the Hakha Lai expression nau ‘the baby’*



In (10), the NP expression *nau* ‘baby’ of type $\langle e, t \rangle$ combines with the iota operator ι (occupying the phonologically null D position) to form the DP expression *nau* ‘the baby’, an expression which can be translated as ‘the unique entity such that said entity is a baby,’ satisfying the conditions for interpreting the bare nominal expression *nau* as a unique immediate situation definite.

To review, Hakha Lai unique expressions surface as bare nominals. Bare nominal NPs receive their definite interpretation via the type shifter iota ι , which yields a definite bare nominal DP which describes the unique entity which satisfies the description of the NP. Next, we will turn

⁵ I assume the DP hypothesis here, though there are some analyses which argue that NPs can exist without D. See Cheng & Sybesma (1999) and Chierchia (1998) for further discussions.

to indexed definites, which contain both the iota operator ι as well as a syntactic representation of the index.

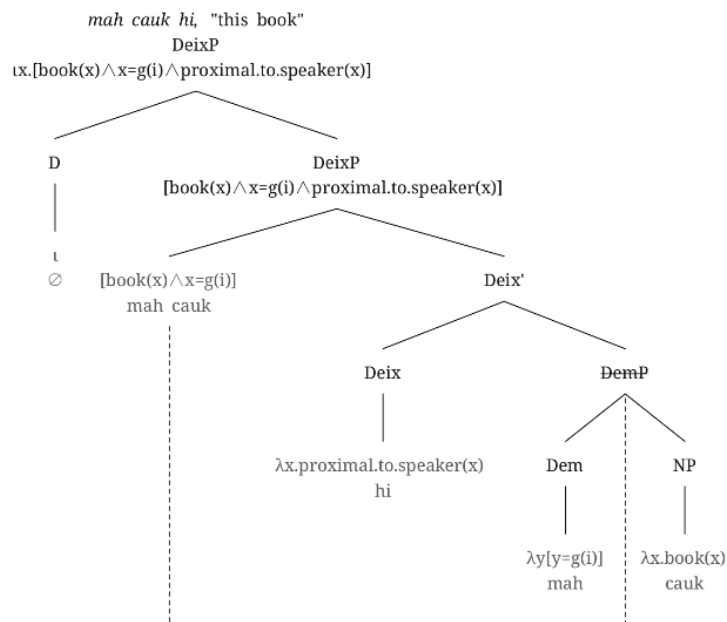
4.2. The semantics of indexed definites

In this section, we will describe the formal semantic structure of indexed definites in Hakha Lai. As a reminder, indexed definites, like unique definites, contain a presupposition of uniqueness. The key way in which they differ is that they contain an additional component: an index. As will be seen in the analysis below, the index is represented morphosyntactically in the structure of the definite expression. Indexed definites in Hakha Lai consist of a D and an index which occupies [Spec, DeixP]. This structure will be articulated further in the sub-sections below. We will begin by looking at demonstratives.

4.2.1. The semantics of demonstratives

Demonstratives in Hakha Lai structurally consist of three critical components. These are the prenominal element *mah*, the head noun (an NP), and the postnominal discourse deictic (which encodes spatial deixis). Before analyzing the semantics, we will look at the syntactic structure of demonstrative expressions in Hakha Lai, using the expression *mah cauk hi* from example (7a). In (11), *mah* is categorized as a Dem head which takes a DP as its argument. The postnominal element, the discourse deictic, is base generated in a superordinate position which I call Deix, the head of a deictic phrase, DeixP. The DemP consisting of *mah* and the NP is raised to [Spec, DeixP].

(11) *The syntactic and semantic structure of the expression mah cauk hi ‘this book’*



In (11), the NP *cauk* ‘book’ is the argument of a demonstrative phrase (DemP) headed by *mah*. The subsequent DemP is the argument of a deictic phrase (DeixP) headed by *hi*, the morphological component which encodes speaker proximal spatial deixis. The DemP is moved to [Spec, DeixP] position to yield the linear order of morphological components found in Hakha Lai demonstrative expressions. The trigger for this movement is not specified here. It is likely a probe from the deictic phrase which seeks a constituent which contains the feature of spatial speaker proximity, though its precise analysis must be left to future research.

Now turning to the semantics of the indexed expression, the NP *cauk* ‘book’ is the argument of the Dem head *mah*, which is equivalent to what Hanink refers to as *idx*. The denotation of *mah* is shown in (12) below.

$$(12) \quad \llbracket mah \rrbracket^g: \lambda y_e [y = g(i)]$$

The prenominal morpheme *mah* combines with the DP *cauk* to yield the DemP *mah cauk*, which contains the properties of being a referent that is assigned an interpretation by the presence of the index and identifies the referent according to an assignment function *g*.

$$(13) \quad \textit{The formal semantics of the demonstrative expression mah cauk}$$

- a. $\llbracket mah \rrbracket^g: \lambda y_e [y = g(i)]$
- b. $\llbracket cauk \rrbracket = \lambda x_{\langle e \rangle}. \text{book}(x)$
- c. $\llbracket mah cauk \rrbracket^g = [\text{book}(x) \wedge x = g(i)]$

Expanding this analysis to include the discourse deictic, *hi* is defined as a spatial deictic morpheme which combines with expressions of type $\langle e \rangle$ to yield another expression of type $\langle e \rangle$ which is interpreted as spatially proximal to the speaker.

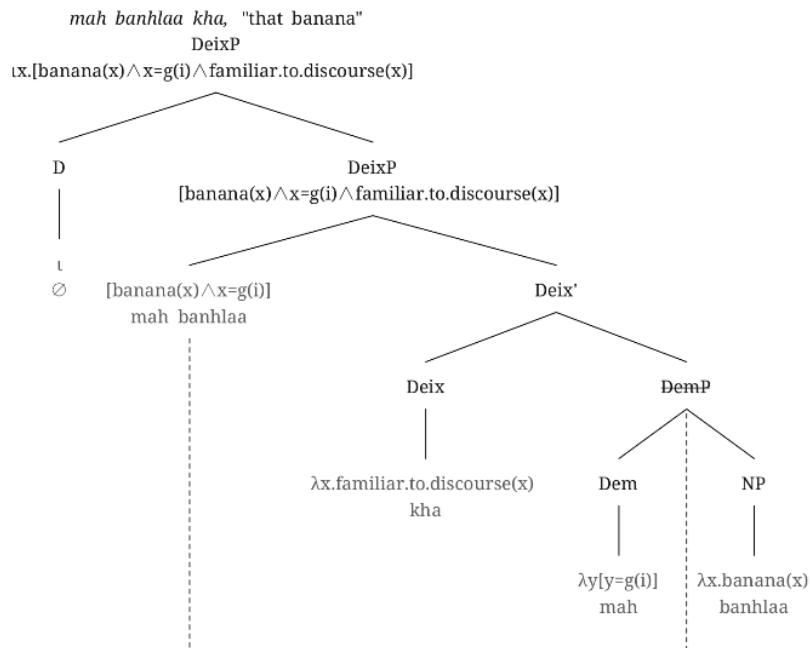
$$(14) \quad \llbracket hi \rrbracket = \lambda x_{\langle e \rangle}. \text{proximal.to.speaker}(x)$$

To review, Hakha Lai demonstratives contain a prenominal element *mah*, which is analyzed as a syntactic instantiation of *idx* (Hanink 2021) and heads a DemP. This prenominal element takes the NP as an argument to form a demonstrative phrase DemP. This DemP is the argument of the postnominal element, a discourse deictic which encodes spatial deictic properties and which heads a deictic phrase, DeixP. The DemP then moves to [Spec, DeixP] position to yield the linear arrangement of morphological components in a demonstrative phrase. DeixP combines with a phonologically null D, which encodes the iota operator ι . The next section offers an analysis of anaphoric definites, a category of indexed definites whose reference is determined by coreference with a prior discourse referent.

4.2.2. The semantics of anaphoric definites

An example of an anaphoric definite is given in example (8) above, *mah banhlaa kha*. Notice that, like demonstratives, this expression contains prenominal *mah* and a postnominal element, *kha*. This example, however, is not an exophoric demonstrative. As seen in example (7) above, postnominal discourse deictics encode spatial deictic properties of the nominal expressions which precede them. In example (8), *kha* does not encode addressee proximity as it usually does and instead is used to mark a familiar referent. Previous studies have shown that *kha* is used in non-spatial deictic contexts to refer to a referent which is familiar to speaker and addressee (Barnes 1998; Wamsley 2023). The specifics of familiar *kha* will not be elaborated on here.

(15) *The syntactic and semantic structure of mah banhlaa kha ‘that banana’*



In (15), *banhlaa* ‘*banana*’, the NP is the complement of *mah*, which is the head of an index phrase *idxP* here called a DemP. Again, *mah* is the index-bearing component in indexed definites. This forms a DemP, which is then the argument of *kha*, the head of a deictic phrase, DeixP. The DemP consisting of *mah banhlaa* is moved to [Spec, DeixP] position to yield the linear order of morphological components in an anaphoric demonstrative expression, *mah banhlaa kha* ‘that banana’.

The difference between the referential function of *mah* in an exophoric demonstrative as in (11) and an anaphoric expression as in (15) is that in an exophoric demonstrative, the assignment function is updated via a gesture towards a referent entity in space (see Ahn 2022 for further discussion). The location in space is further contextualized by the postnominal discourse deictic. In an anaphoric expression, a referent is introduced into the discourse by first mention, thus

updating the common ground and providing a referent to be assigned an identity via the assignment function. The semantic composition of *mah banhlāa* is shown below in (16).

(16) *The formal semantics of the expression mah banhlāa*

- a. $[[mah]]^g: \lambda y_e [y = g(i)]$
- b. $[[banhlāa]] = \lambda x_{(e)}. banana(x)$
- c. $[[mah banhlāa]]^g = [banana(x) \wedge x = g(i)]$

Like with exophoric demonstratives, the pronominal and nominal elements *mah banhlāa* combine with a postnominal discourse deictic. In example (8), this is *kha*. The denotation of *kha* is shown in (17) below.

(17) $[[kha]] = \lambda x_{(e)}. familiar.to.discourse(x)$

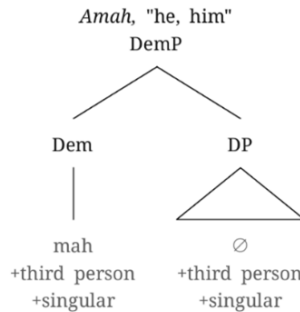
Unlike *hi* in the example above, *kha* encodes familiarity to speaker and addressee, as opposed to a spatial deictic property.

To review, anaphoric expressions are similar in structure to exophoric demonstratives. The difference is where they receive their reference from (pointing vs. previous mention in discourse). Anaphoric expressions in Hakha Lai consist of a pronominal element, *mah* which is the head of *idxP*, the nominal, and a postnominal discourse deictic, which in this example, is *kha*, marking speaker and addressee familiarity. The next section analyzes Hakha Lai personal pronouns. The analysis here treats pronouns as another kind of indexed definite expression wherein *mah* is a suffixal as opposed to pronominal element.

4.3. The semantics of personal pronouns

Personal pronouns in Hakha Lai come in two forms, a “focus” and a “contrast form, to use Hlun (2007)’s terminology. This study will only examine the focus pronouns, leaving contrast pronouns to future research. Focus pronouns in Hakha Lai are composed of a suffixal base *mah* and a prefixal element which encodes person and number e.g., *keimah* ‘I’, *kanmah* ‘we’, *nangmah* ‘you’, *nanmah* ‘you (pl)’, etc. The syntactic and semantic analysis of Hakha Lai personal pronouns described here is based on Elbourne’s (2005) notion that personal pronouns are definite descriptions. As such, the function of the *mah* suffix in each of the pronouns is the same as its function in previously described indexed definite expressions, the index-bearing *idx* head. In fact, as the analysis below shows, it is the index head itself which is the personal pronoun. The surface representation of personal pronouns which encode phi-features person and number are merely surface-level properties of *mah*.

The morphosyntax of *amah* ‘him’ from the sentence in (9), is shown in (18) below.

(18) *Morphosyntax of amah 'him'*

As shown in (18), *mah*, the head of a DemP, takes an elided DP as an argument and surfaces with the phi-features of the elided DP. In (18), this is third person and singular and thus, the pronominal element *mah* surfaces with pronominal *a*- denoting third person singular reference. The semantic structure of personal pronouns is shown below in (19).

(19) $[[\text{-mah}]^g: \lambda y_e [y = g(i)]$

Once again, the role of *mah* is to take an entity argument and denote that the entity's reference is determined by the assignment function. Like anaphoric definites, the reference from the assignment function is a result of the referent having been mentioned previously in discourse.

This ends the syntactic and semantic analysis of Hakha Lai personal pronouns and the section on indexed definites in Hakha Lai. All three categories of indexed definites described here contain a morphological element, *mah*. In exophoric demonstratives and anaphoric definites, this element appears in pronominal position and identifies the reference of the entity described in the definite expression. In the case of personal pronouns, this morphological element surfaces with phi-features that match the referent. The next section is a discussion of how the analysis of Hakha Lai definite expressions contributes to ongoing research on different categories of definites across language. This discussion is followed by some concluding remarks.

5. Discussion

This paper has discussed the syntactic and semantic structure of definite expressions in Hakha Lai, focusing on two categories, "unique" definites and "indexed" definites. The difference between these two categories as described in previous research is the presence of an index in indexed definites, meaning that the reference of an indexed definite entity is given via an assignment function in addition to the presupposition of uniqueness. This paper looked at both categories of definite expressions, including three sub-categories of indexed definites. Just as indexed definites differ from unique definites in that they have an additional semantic structural property, the morphological structure of indexed definites differs as well, namely in the presence of a morpheme *mah*. This morpheme has been analyzed here as *idx*, a syntactic instantiation of the index, a crucial component in the interpretation of indexed definites. This paper has looked at the presence of the

index in exophoric demonstratives, anaphoric definite descriptions, and personal pronouns, all expressions whose reference is obtained through an index.

6. Conclusion

This paper has investigated two types of definite expressions in natural language, unique definites and indexed definites. Data from speakers of Hakha Lai, a South Central Tibeto-Burman language spoken in northwest Burma and by a community of speakers in Indianapolis, was used to analyze these two categories of expressions. The key findings of the analysis are that while both types of definite expression contain a presupposition of uniqueness, indexed definites, including exophoric demonstratives, anaphoric definites, and personal pronouns contain an additional semantic component in the form of an index which assigns reference via an assignment function. The key contribution of this study is that the index is represented in the syntactic structure of indexed definites through an overt morphological marker, *mah*. Previous research has investigated the syntactic representation of such markers (cf. Hanink 2021) and this research continues this line of inquiry. As the syntax and semantics of such expressions are investigated further, we will learn more about the nature of the syntactic representation of indices in definite expressions and elsewhere in natural languages.

References

- Ahn, D. (2022). Indirectly direct: An account of demonstratives and pointing. *Linguistics and Philosophy* 45. 1345-1393.
- Arkoh, R. & L. Matthewson. (2013) A familiar definite article in Akan. *Lingua* 123 (2013). 1-30.
- Baclawski, K. (2013) The Kuki-Chin deictic complex: examining interconnected reference systems. *Linguistic Society of America 87th Annual Meeting*. Presentation Slides. https://linguistics.berkeley.edu/~kbaclawski/Baclawski_LSA_2013presentation.pdf (15 September 2023)
- Barnes, J. (1998) “Tsuu khaa tii hla?: Deixis, Demonstratives and Discourse Particles in Lai Chin,” *Linguistics of the Tibeto-Burman Area* 21(1). 53–86.
- Berkson, K, J. Wamsley, S. Lotven, S. Chelliah, K. Van Bik, S. Champlin, K. Sakhong, S. H. Par, A. Matthews & A. Bohnert. (2023). A developing community of collaboration in Indiana. In B. Murray, M. Brill-Carlat & M. Höhn, (eds.) *Migration, Displacement, and Higher Education: Now What?* (Political Pedagogies). Palgrave Macmillan. 23-35.
- Cheng, L. & R. Sybesma. (1999) Bare and not-so-bare nouns and the structure of NP. *Linguistic Inquiry* 30(4). 509-542.

- Chierchia, G. (1998) Reference to kind across languages. *Natural Language Semantics* 6. 339-405.
- Dayal, V. (1999). Bare NP's, reference to kinds, and incorporation. In Matthews, T. and D. Strolovitch (eds.) *Proceedings of SALT IX*. 34-51.
- Elbourne, P. (2005). *Situations and individuals* (Vol. 90). Cambridge, MA: MIT press.
- Hanink, E. (2021) DP Structure and internally headed relatives in Washo. *Natural Language & Linguistic Theory* 39. 505-554.
- Hawkins, J. (1978) *Definiteness and indefiniteness*. London: Croom Helm.
- Heim, I. (1982) *The Semantics of definite and indefinite noun phrases*. Amherst, MA: University of Massachusetts Ph.D. dissertation.
- Hlun, C. (2007) Pragmatic influence on pronouns in Lai (Hakha) Chin, with especial reference to focus and contrast. In Wayland, R., Hartmann, J. & Sidwell, P. (eds.) *SEALSXII: Papers from the 12th meeting of the Southeast Asian Linguistics Society (2002)*. 79-88. Canberra, Australia: Pacific Linguistics.
- Jenks, P. (2018) Articulated definiteness without articles. *Linguistic Inquiry* 49(3). 501-536.
- Jenks, P., & R. Konate (2022). Indexed definiteness. *Glossa: a journal of general linguistics*, 7(1).
- Moroney, M. (2021). Updating the typology of definiteness: Evidence from bare nouns in Shan. *Glossa: a journal of general linguistics* 6(1): 58. 1-28.
- Salaz, S. & S. Raymer. (2020, December 12) "Welcome to Chindianapolis." Indianapolis Monthly. <https://www.indianapolismonthly.com/arts-and-culture/circle-city/welcome-to-chindianapolis> (Accessed 21 December 2023).
- Schwarz, F. (2009) *Two types of definites in natural language*. Amherst, MA: University of Massachusetts Ph.D. dissertation. 1-316.
- Wamsley, J. (2023) *Topics in Hakha Lai Nominal Marking*. Bloomington, IN: Indiana University Ph.D. dissertation. 1-285.

Inclusive plural in a ‘general number’ language¹

Virginia Dawson — Western Washington University

Abstract. This paper examines plural marking in the general number language Tiwa (Tibeto-Burman; India). I show that Tiwa has an additive plural marker that’s fully productive, though optional. I also show that despite what might be an initial appealing analysis of Tiwa’s plural marker as semantically exclusive (in contrast to Tiwa’s inclusive bare nouns), Tiwa’s plurals receive clearly inclusive readings just like English plurals do. The key typological finding of this work is that general number languages can have inclusive plural markers. I also discuss implications for how plurality is analyzed, arguing that a traditional quantity implicature approach to exclusive readings cannot be maintained for the Tiwa data.

1 Introduction

There are long-standing debates in the semantic literature about the proper treatment of plural markers. One focus of these debates is on how to capture the generalization that plural marking in languages like English yields ‘exclusive’ readings in certain contexts and ‘inclusive’ readings in others. That is, while the plural marked noun *children* in (1a) is interpreted roughly as ‘more than one child’ (**excluding** atomic children), the same form in (1b) is usually interpreted as ‘at least one child’ (**including** atomic children).

- | | | | |
|-----|----|---|------------------|
| (1) | a. | Alex has children. | EXCLUSIVE PLURAL |
| | | ≈ Alex has more than one child. (no atoms; only sums) | |
| | b. | Does Alex have children? | INCLUSIVE PLURAL |
| | | ≈ Does Alex have one or more children? (atoms and sums) | |

There are several distinct proposals to deal with the two readings that plural marking in English and other languages receives. Sauerland et al. (2005), for example, propose that plural marking is semantically inclusive, but exclusive readings arise in non-downward-entailing environments, such as (1a), due to competition with the atomic singular form. Farkas and de Swart (2010), in contrast, propose that plural marking is polysemous between an inclusive and exclusive reading, and that a pragmatic principle ensures that the strongest possible meaning is chosen in a given context.

Most of the debate has centered on languages with an obligatory singular-plural distinction, such as English and Hungarian. While these languages show differences in whether plural marking is required or allowed in particular morphosyntactic contexts, they are similar in having a core singular-plural distinction. More recently, data from a wider array of languages has come into consideration. Martí (2020), for example, focuses on languages with dual and other number categories, arguing that an ambiguity/polysemy approach is better able to handle such systems than the implicature approach proposed by Sauerland et al. (2005).

¹Thanks to my Tiwa teachers Bibiana Maslai, Mary Maslai, Olphindro Malang, and the late Juliana Maslai, for sharing their language with me. Thanks also to the participants of Triple A10, who provided incredibly useful questions, comments, and feedback. Any errors are entirely my own.

In this paper, I introduce a different kind of data from Tiwa, a Tibeto-Burman language of northeast India, which have clear consequences for theories of the semantics of plural marking. Tiwa, like many other languages of the world, is a ‘general number’ language (Corbett 2000), meaning that the domain of unmarked nouns contains both atomic and plural individuals. For example, the sentence in (2), which contains the unmarked noun *miyâw* ‘cat’, is number neutral: it is true whether Mukton saw one cat or more than one cat. This stands in stark contrast to its possible English translations, which necessarily show a singular/plural distinction on the noun.²

- (2) Mukton **miyâw**-go nú-ga.
 Mukton cat-ACC see-PFV
 ‘Mukton saw a cat/cats.’ [MM 2017.1.46]

It has sometimes been assumed in the formal semantics literature that such languages do not have productive plural marking (e.g. Chierchia 1998). Where ‘plural’ markers do exist in such languages, they are often highly restricted (e.g. to human-denoting nouns), encode other semantic information (notably definiteness), and additionally have ‘associative’ uses.³ There are general number languages with dedicated additive plural marking, however, and Tiwa is one of them. In particular, Tiwa has a dedicated and fully productive (though optional) plural marker *-râw*, illustrated in (3). In contrast to (2), the sentence in (3) must convey that Mukton saw more than one cat, just like its English translation.

- (3) Mukton **miyâw-raw**-go nú-ga.
 Mukton cat-PL-ACC see-PFV
 ‘Mukton saw cats.’ [MM 2017.1.46]

Given the possible interpretations of (2) and (3), a reasonable assumption about the semantics of number in a language like Tiwa could be that bare nouns have an inclusive semantics (ranging over both atoms and sums), while plural-marked nouns have an exclusive semantics (ranging only over sums). Indeed, this assumption has been made for bare nouns and *tul*-marked nouns in Korean, another general number language with optional plural marking (Kang 1994; Rullmann and You 2006). However, the main contribution of this paper is to show that such an assumption cannot be made, at least for Tiwa. I show that plural-marked nouns in Tiwa can receive inclusive readings in just the same way that English plural-marked nouns can. That is, general number languages can have inclusive plurals.

This finding is an important typological point, but also has theoretical consequences. In particular, the competition account of the exclusive/inclusive plural alternation proposed by Sauerland et al. (2005) cannot account for the Tiwa data, as there is no atomic form in the language to compete with the plural to give rise to exclusivity inferences. In contrast, the absence of an atomic form does not cause problems for ambiguity/polysemy approaches. The existence of inclusive plurals in Tiwa also raises questions about variation in the semantics of (dedicated additive) plural markers across languages. Specifically, since inclusive plurals can exist in a language with number-neutral bare nouns,

²Tiwa examples are presented in the orthography of Joseph’s (2014) dictionary. A reference to speaker, year, notebook number and page number are given after each translation. Abbreviations are: ACC ‘accusative’, AUX ‘auxiliary’, CL ‘classifier’, COMP ‘complementizer’, COP ‘copula’, FOC ‘focus’, GEN ‘genitive’, INTS ‘intensifier’, IPFV ‘imperfective’, LOC ‘locative’, NEG ‘negation’, NEUT ‘neutral aspect’, NMLZ ‘nominalizer’, PFV ‘perfective’, PL ‘plural’, SG ‘singular’.

³Such plural markers include Mandarin Chinese *-men* (Li 1999) and Japanese *-tachi* (Nakanishi and Tomioka 2004).

it raises the question of whether **any** language has plural-marking that is purely exclusive. Recent empirical work by Renans et al. (2020) lends support to this possibility, showing experimentally that Turkish plurals can receive inclusive readings, despite claims in the literature to the contrary. More careful empirical work is needed to test whether this holds up cross-linguistically.

The remainder of this paper is structured as follows. In Section 2, I introduce the plural marker *-râw*, showing that it is a dedicated and productive (though optional) plural marker. In Section 3, I show that despite what we might initially expect for a general number language, plural-marked nouns in Tiwa have inclusive readings. I discuss some typological and theoretical consequences that this finding has in Section 4, and compare the Tiwa case to a similar but crucially different system in Buriat (Bylinina and Podobryaev 2020). I conclude in Section 5.

2 Plural marking in Tiwa

Tiwa is a Tibeto-Burman language of the Boro-Garo subgroup spoken by approximately 33,900 people primarily in Assam, northeast India (Post and Burling 2017; Eberhard et al. 2023).⁴ I collected the data presented here in Umswai, Karbi Anglong district, Assam, between 2015 and 2023, through work with three speakers. The data were collected primarily through elicitation, following the methodology for semantic fieldwork laid out in Matthewson (2004). For more information on Tiwa, including a basic grammar sketch, see Dawson (2020). In this section, I lay out the properties of the plural marker *-râw*, showing that it is a dedicated and productive additive plural marker, and discussing its distribution, including its optionality.

2.1 *-râw* is a dedicated productive plural marker

Plural markers in many general number languages are restricted in their distribution, and are not strictly additive plural markers. Mandarin *-men*, for example, only appears on pronouns and human nouns, and encodes definiteness in addition to plurality (Li 1999). Similarly, Japanese *tachi* shows properties of being an associative plural marker, meaning that it combines with a referential expression and returns a plurality consisting of that individual and other individuals associated with it (Nakanishi and Tomioka 2004). Like Mandarin and Japanese, Tiwa has these sorts of plural markers: it has both a plural marker *-mân*, which encodes definiteness, and *chógol*, which is an associative plural marker (Dawson 2020).

In contrast, *-râw* is a fully productive and purely additive plural marker. Unlike Mandarin *-men*, *-râw* can appear on any count noun. This includes nouns denoting humans, as shown in (4) with the noun *loró* ‘priest’, animate non-humans, as shown in (3) above with the noun *miyâw* ‘cat’, and inanimate objects, as shown in (5) with the noun *táp* ‘penknife’. It can also appear on countable abstract nouns, such as *kit* ‘song’, as shown in (6).⁵

⁴This population estimate comes from the 2011 census, as reported in Eberhard et al. (2023).

⁵In contrast, *-râw* cannot ordinarily occur on mass nouns (unless they are coerced). This applies to a wide range of substance-denoting nouns, but is illustrated here with *kakhîr* ‘milk’:

- (1) * Mansing **kakhîr-raw**-go nung kar-ga.
 Mansing milk-PL-ACC drink AUX-PFV
 Intended: ‘Mansing drank up the milk(s).’

[MM 2018.1.130]

- (4) Mukton loró-**râw**-go nú-ga.
Mukton priest-PL-ACC see-PFV
'Mukton saw (the) priests.' [MM 2017.1.48]
- (5) Ang táp-**râw**-go mokhói hál-ga.
1 SG penknife-PL-ACC drop AUX-PFV
'I dropped (the) penknives.' [BM 2017.2.36]
- (6) Ang kit-**raw**-go rojá-ga.
1 SG song-PL-ACC sing-PFV
'I sang (the) songs.' [BM 2018.1.148]

Also in contrast to Mandarin *-men* (and Tiwa *-mân*), *-râw* does not encode definiteness. There are several pieces of evidence that *-râw*-marked nouns can be indefinite.⁶ First, *-râw* plurals can be used to make existential claims, as shown in (7) and (8).

- (7) Nó níng-o khódo-**râw** tong-o.
house inside-LOC mosquito-PL exist-NEUT
'There are mosquitoes inside the house.' [MM 2017.1.65]
- (8) Australia-w tamúr-lô kanggaro-**raw** tong-o.
Australia-LOC many-FOC kangaroo-PL exist-NEUT
'There are many kangaroos in Australia.' [BM 2018.2.8]

Similarly, *-râw* plurals can be used to introduce new discourse referents. This is illustrated in (9), which shows *hâdi-raw* 'elephant-PL' and *makhri-râw* 'monkey-PL' introducing new discourse referents in a text.

- (9) Preceding discourse [in Tiwa]: Once, Saldi was working in the paddy field.
Ashôbai-lo, krom-e phána **hâdi-raw** arô **makhri-râw** cholói phi-ga.
suddenly-FOC forest-GEN from elephant-PL and monkey-PL run AUX-PFV
'Suddenly, elephants and monkeys ran out of the forest.' [MM 2017.1.63]

Unlike definite plural DPs in languages like English, *-râw* plurals in Tiwa do not come with maximality requirements. For example, the plural-marked noun *lái-râw* 'book-PL' in sentence in (10) can be used in a context where the proposition is true only of some of the books in the context, and false of others.

- (10) Context: There are ten books on the shelf, and nothing else. Five of the books fell down.
✓ **Lái-râw** kói phi-ga.
book-PL fall AUX-PFV
'Books fell.' [MM 2018.2.30]

Finally, *-râw* plurals can also serve as the predicate of a sentence. This is shown in (11), which conveys simply that Monbor and Mukton have the property of being priests (rather than identifying them as a particular pair of priests).

⁶See Dawson and Jenks (2023) for a summary of diagnostics for (in)definiteness.

- (11) Monbor re Mukton **loró-râw** hóng-do.
 Monbor and Mukton priest-PL COP-IPFV
 ‘Monbor and Mukton are priests.’ [MM 2017.1.52]

Note that while *-râw* does not encode definiteness, as the examples above show, *-râw*-marked plurals can have a definite interpretation in argument position in the same way that bare non-plural-marked nouns in Tiwa can (Dawson 2020). What is important for our purposes is that the plural marking itself does not contribute definiteness.

-râw plurals in Tiwa also cannot receive associative plural readings. First, *-râw* cannot combine with a referential expression such as a proper name, as shown in (12).⁷ Similarly, when it combines with a noun like *rajâ* ‘king’, as in (13), it can only receive an additive plural reading, rather than an associative reading. (In contrast, Tiwa’s associative plural marker *chógol* can be used with proper names and with nouns to create associative meanings; Dawson 2020.)

- (12) * Lastoi-**raw** phi-ga.
 Lastoi-PL come-PFV
 Intended: ‘Lastoi and company came.’ [BM 2018.2.9]
- (13) Rajâ-**raw** phi-ga.
 king-PL come-PFV
 ‘The kings came.’ [BM 2018.2.10]
 Cannot mean: ‘The king and his associates came.’

The data presented in this section show that *-râw* is a dedicated, additive plural marker, similar to plural marking in English and other non-general number languages.

2.2 *-râw* is freely available and always optional

While *-râw* has a similar semantics to plural-marking in languages like English and Hungarian, its distribution is different. In particular, *-râw* is (i) always available when the domain contains pluralities, and (ii) always optional.

Both these properties are evident for bare nouns in argument position. Plural marking is always available when the domain of the (count) noun contains pluralities, regardless of animacy properties of the noun, its grammatical role, and its case marking. Plural marking is also optional for bare nouns in argument position, although it is preferred for human and highly animate nouns in most contexts.⁸ This was shown for sentences (2) and (3) above, where plural-marking is not required when the domain contains pluralities. Examples (14)-(16) provide further examples of unmarked nouns being used when the domain contains pluralities.

⁷*-râw* is also not found in Tiwa’s plural pronouns: *ching* ‘1PL’, *nábur* ‘2PL’, and *pibúr* ‘3PL’. This again contrasts with the behavior of Chinese *-men* and Japanese *tachi*.

⁸Note that by “optional” I mean that it is never required to yield a grammatical sentence and a sentence without plural marking can be used truthfully in the same contexts that the plural-marked alternative could be used. While there are some pragmatic differences, these are often difficult to detect, and speakers frequently comment that sentences with *-râw* and without *-râw* are interchangeable, except where plural marking is preferred for human nouns and dispreferred for inanimate nouns in some contexts. Further work is required to explore these pragmatic effects; see §4.3 for discussion.

- (14) Hé-w **miyâw** parâ tong-o, arô hé-w **miyâw** khom tong-o.
 here-LOC cat more exist-NEUT and here-LOC cat fewer exist-NEUT
 ‘There are more cats here, and there are fewer cats here.’ [pointing] [BM 2018.3.68]
- (15) **Khódo** khúp phi-do.
 mosquito INTS come-IPFV
 ‘(A) mosquito(es) keep coming.’ [OM 2022.1.79]
 (Explicitly judged by the consultant to convey that either one or more than one mosquito kept coming.)
- (16) [Lastoi-ne táw-a **kashóng** kojá-gô] Saldi pre-ga.
 Lastoi-GEN weave-NMLZ dress red-ACC Saldi buy-PFV
 ‘Saldi bought the dress(es) that Lastoi wove.’ [BM 2017.2.48]
 (Offered as a translation into Tiwa of the plural English sentence.)

The availability and optionality of plural marking extends to nouns that are modified by numerals and quantifiers that entail plurality. For example, plural marking is always available, but never required, for numerals greater than ‘one’. This is illustrated for the numeral *kiníng* ‘two’ in (17), which shows that the noun *miyâw* can be marked with *-râw* or left unmarked.

- (17) Mukton [**kiníng** miyâw(-**raw**)-go] nú-ga.
 Mukton two.CL cat(-PL)-ACC see-PFV
 ‘Mukton saw two cats.’ [MM 2017.1.50]

Similarly, quantifiers like *mile* ‘all’ and *pángai* ‘many’ also appear with plural-marked nouns and un-plural-marked nouns, as shown in (18) and (19)-(20).

- (18) Mukton [**mile** miyâw(-**raw**)-go] nú-ga.
 Mukton all cat(-PL)-ACC see-PFV
 ‘Mukton saw all the cats.’ [MM 2017.1.50]
- (19) Ang [**pángai** korkhyá-**râw**-go] lak mán-ga.
 1SG many child-PL-ACC meet-PFV
 ‘I met many children.’ [BM 2015.1.97]
- (20) Mukton [**pángai** libíng-gô] marê-ga.
 Mukton many person-ACC kill-PFV
 ‘Mukton killed many people.’ [MM 2017.2.29]

Finally, plural-marking is also optional for predicative nouns. Sentence (11) above shows that a *-râw*-marked noun can serve as the predicate of the sentence; sentence (21) below shows that a bare noun can also be predicated of a plurality.

- (21) Monbor re Mukton **loró** hóng-do.
 Monbor and Mukton priest COP-IPFV
 ‘Monbor and Mukton are priests.’ [MM 2017.1.52]

3 Inclusive plurality and general number

In the previous section I laid out the basic properties of *-râw* plurals in Tiwa. I showed that *-râw* is a dedicated additive plural marker, which is fully productive, though optional. This optionality is, of course, tied to Tiwa's status as a general number language: bare nouns are number neutral and can thus range over domains that include pluralities. In this section, we consider how plural marking in a general number language might be modeled as purely exclusive, given the data above. I then present evidence that *-râw* plurals can receive inclusive readings, just as plurals in English can.

3.1 Are plurals purely exclusive?

Bare nouns in general number languages are number neutral. A reasonable assumption about the meaning of plural-marked nouns in such languages is that they, in contrast, are not number neutral; instead, their domains contain only pluralities. On this view, a language like Tiwa has the number contrast represented in (22). In such languages, the extension of bare nouns like *miyâw* 'cat' is the set of atomic cats and sums of cats, while the extension of plural marked nouns like *miyâw-râw* 'cats' is the set that contains only sums of cats.

- (22) a. $\llbracket \text{miyâw} \rrbracket = \lambda x. *cat(x)$ Bare noun
 = $\{a, b, c, a+b, a+c, b+c, a+b+c, \dots\}$
 b. $\llbracket \text{miyâw-râw} \rrbracket = \lambda x. *cat(x) \ \& \ SUM(x)$ Plural-marked noun
 = $\{a+b, a+c, b+c, a+b+c, \dots\}$

This is the analysis proposed by Kang (1994) for Korean, which is also a general number language with plural marking, as well as Rullmann and You (2006). It is also the analysis I assumed in my earliest work on plurality in Tiwa (Dawson 2018).

This view is appealing because it provides a straight-forward analysis of the data presented above, capturing the difference in semantics between bare nouns and plural-marked nouns in general number languages. For example, the analysis in (22) yields the following truth conditions for sentences (2) and (3) above respectively: (2) is true in scenarios where Mukton saw one cat or more than one cat, while (3) is only true if Mukton saw more than one cat.

- (23) a. $\llbracket (2) \rrbracket = \exists x[*cat(x) \ \& \ saw(\text{Mukton}, x)]$
 b. $\llbracket (3) \rrbracket = \exists x[*cat(x) \ \& \ SUM(x) \ \& \ saw(\text{Mukton}, x)]$

If this analysis is correct, it also suggests an interesting point of cross-linguistic variation in the semantics of plural-marking across languages. Specifically, where the domain of a plural-marked noun in a language like English must (at least sometimes) contain atoms, the domain of a plural-marked noun in a general number language would never. Plural marking in the two sorts of languages would also have a different compositional contribution: in a language like Tiwa, plural marking is subtractive, removing atoms. In contrast, in a language like English – on the assumption that bare nouns are atomic – plural marking is additive, with sums being added to the domain of the noun.

3.2 Plurals in Tiwa can be inclusive

While an analysis of plural marking in Tiwa as semantically exclusive is initially appealing, it is unable to account for the full range of data. Specifically, *-râw*-marked plurals in Tiwa can receive inclusive readings in just the same way that plural-marked nouns in languages like English can. If the domain of plural-marked nouns contains only sums, inclusive readings should be unavailable. This is simply not borne out by the data.

Inclusive readings can be easily seen in the behavior of indefinite *-râw* plurals in questions, as shown in examples (24)-(26), which report the clear-cut judgments of three distinct speakers. (24) shows a polar question that contains *korkhyá-râw* ‘child-PL’, which is a natural way to ask the addressee if they have children (similar to the English translation given). When asked how the addressee would respond if they had only one child, the consultant judged that the addressee would respond in the affirmative, as in (24a), and that it would be infelicitous to respond in the negative, as in (24b). This indicates that the domain of *korkhyá-râw* contains atomic children.

- (24) Né korkhyá-**râw** tong-o ná?
 your child-PL exist-NEUT PQ
 ‘Do you have children?’
- a. Oi, ái sája korkhyá tong-o.
 yes my one.CL child exist-NEUT
 ‘Yes, I have one child.’
- b. # Cha, ái sája-side korkhyá tong-o.
 no my one.CL-only child exist-NEUT
 ‘No, I have only one child.’ [OM 2022.1.22-24]

The data in (25) show a similar result from another speaker. Here the polar question contains *makhrí-râw* ‘monkey-PL’. This speaker also judged it felicitous to respond in the affirmative even if the addressee saw only one monkey.

- (25) Ná makhrí-**râw**-go nú-ga ná?
 you monkey-PL-ACC see-PFV PQ
 ‘Did you see monkeys?’
- a. Oi, kishá makhrí-gô nú-ga-ng.
 yes one.CL monkey-ACC see-PFV
 ‘Yes, I saw one monkey.’ [BM 2018.2.11]

The data in (26) show the same question, but with judgments on answers from another speaker. Here, the speaker rejected a negative answer in the case that the addressee saw only one monkey (26a), saying instead that a negative answer would require that the addressee had seen no monkeys, as in (26b). Like the other two speakers, this speaker accepted an affirmative answer when the proposition is true of only one atomic individual, as shown in (26c).

- (26) Ná makhrí-**râw**-go nú-ga ná?
 you monkey-PL-ACC see-PFV PQ
 ‘Did you see monkeys?’

- a. # Cha, kishá-sidê makhrí-gô nú-ga-ng.
no one.CL-only monkey-ACC see-PFV
'No, I only saw one.'
- b. Cha, kishá-gô-bo nú-ya-m.
no one.CL-ACC-even see-NEG-PST
'No, I didn't see any.'
- c. Oi, thêbo kishá-sidê makhrí-gô nú-ga-ng.
yes but one.CL-only monkey-ACC see-PFV-1SG
'Yes, but I only saw one.'
- [MM 2018.1.118]

Clear inclusive readings of *-râw*-marked plurals also emerge under negation. For example, the sentence in (27) conveys that there are no kangaroos at all in Umswai; it does not convey that there are no pluralities of kangaroos (while there may be a single one). Atoms must be in the domain of the plural marked noun to yield this reading.

- (27) Umswai-o kangaro-**raw** cha.
Umswai-LOC kangaroo-PL NEG.EXIST
'There are no kangaroos in Umswai'
- [BM 2018.2.8]

A similar example is given in (28). This is a natural sentence that conveys that Monbor doesn't like children in general. It does not convey that Monbor doesn't like groups of children (but that he is okay with singular children).

- (28) Monbor korkhyá-**râw** kumún mán-ya.
Monbor child-PL good get-NEG
'Monbor doesn't like children.'
- [MM 2023.1.53]

Both the question data and the negation data show that *-râw* plurals in Tiwa have inclusive readings, in just the same way that English plurals do, showing that an exclusive-plural approach as sketched out in §3.1 cannot be adopted for Tiwa. These readings emerge despite the fact that a number-neutral bare noun can also be used in all the above sentences to the same effect. That is, any of the questions in (24)-(26) above can be asked with a bare noun in place of the plural-marked noun. Similarly, the sentences in (27) and (28) can be expressed with bare nouns to the same effect.

4 Broader implications

The data presented in the previous section shows that *-râw*-marked plurals in Tiwa can and do receive inclusive readings. This finding has broader implications, both typological and theoretical, which will be discussed in this section. The section concludes with a preliminary discussion of how to analyze the Tiwa data.

4.1 General number languages can have inclusive plurals

The Tiwa data clearly illustrate that general number languages can have inclusive plurals. This is an important finding, because it goes against an (often implicit) assumption in the literature that languages with number neutral nouns only have exclusive plurals (e.g. Rullmann and You 2006,

Bale et al. 2011). As discussed in §3.1, this generalization is appealing in that it easily differentiates the meaning of the plural-marked form from the bare form of the noun. As the data in §3.2 show, however, that assumption (at least for Tiwa) is not supported by the data.

At this stage, it remains an open question whether all plurals in general number languages are inclusive, or whether plurals in some languages are indeed semantically exclusive. There are at least two other general number languages that I'm aware of that have inclusive plurals: Ch'ol, a Mayan language (Carol Rose Little p.c.), and Buriat, a Mongolic language (Bylinina and Podobryaev 2020). (Buriat will be discussed in more detail in §4.2 below.) The presence of inclusive plurals in unrelated general number languages in different parts of the world suggests that they may be relatively widespread.

While several languages discussed in the literature have been claimed to have semantically exclusive plurals, it's unclear to which extent this holds up under rigorous empirical testing. Turkish, for instance, has been claimed to have exclusive plurals (e.g. Bale et al. 2011), but recent experimental work by Renans et al. (2020) has shown that plurals in Turkish can and do receive inclusive readings in some contexts. Further testing is needed to establish whether exclusive analyses of plural markers in other general number languages, such as Korean, need to be revised.

4.2 Exclusive plurality is not (always) a quantity implicature

In addition to the typological finding discussed above, the Tiwa data have implications for semantic theories of plural marking. In particular, the Tiwa data cannot be captured by theories that derive exclusive readings of plural as implicated via competition with an atomic alternative.

The data above show that *-râw* plurals in Tiwa, like English plurals, receive both inclusive and exclusive readings, depending on the context. An adequate analysis of Tiwa plurals thus needs to derive both readings in their appropriate contexts. One influential approach is to derive the exclusive reading of plurals as a quantity implicature (Sauerland et al. 2005). On such accounts, plurals are assumed to be semantically inclusive; their domain includes both atoms and sums, as in (29a). In contrast, singular forms are atomic; their domain does not include sums, but only atoms, as shown in (29b).

- (29) a. $\llbracket \text{noun-PL} \rrbracket = \lambda x. * \text{noun}(x)$
 b. $\llbracket \text{noun-SG} \rrbracket = \lambda x. * \text{noun}(x) \ \& \ \text{ATOM}(x)$

These forms are in competition, and since (29b) entails (29a) in upward-entailing environments, a quantity implicature is triggered. (This analysis can be implemented in various ways; see Sauerland et al. 2005, Spector 2007, and Zweig 2009. I set these differences aside here, since they do not affect whether the approach can apply to Tiwa or not.)

An implicature approach of this nature cannot be responsible for exclusive readings of plural in Tiwa because there is no atomic form to act as competitor to the plural. As was shown above, unmarked nouns in Tiwa are number neutral; their domains can include sums. There are no other plausible candidates for an atomic competitor for the plural-marked noun on standard approaches to determining alternatives. Nouns modified by 'one', for example, are too structurally distinct on most approaches to quantity implicatures (Katzir 2007; Fox and Katzir 2011).⁹

⁹This argument relies on the assumption that the alternatives relevant for calculating implicatures must correspond

At this point, it is worth comparing the Tiwa data to Buriat, another general number language with inclusive plurals. Bylinina and Podobryaev (2020) argue for a quantity implicature account of exclusive readings of plural-marked nouns in Buriat, along the lines discussed above. In particular, they argue that apparently unmarked nouns in Buriat are actually ambiguous between a form that lacks a NumP projection and one that has a NumP projection with a null singular morpheme. Bare nouns that lack a NumP projection are number neutral, as in (30a). Plural-marked nouns are similarly number neutral (i.e. inclusive), as in (30b). In contrast, apparently bare nouns that have a NumP projection with a null singular morpheme are atomic, as in (30c).

- (30) a. $\llbracket \text{noun} \rrbracket = \lambda x. * \text{noun}(x)$
 b. $\llbracket \text{noun-PL} \rrbracket = \lambda x. * \text{noun}(x)$
 c. $\llbracket \text{noun-SG} \rrbracket = \lambda x. * \text{noun}(x) \ \& \ \text{ATOM}(x)$

Bylinina and Podobryaev argue that exclusivity implicatures arise for plural-marked nouns in Buriat because they are in competition with the null singular form. In contrast, exclusivity implicatures do not arise for bare nouns that lack a NumP projection, despite having the same semantics as plural nouns, because the null singular form is too structurally complex to compete with it.

Crucially, evidence for the existence of the null singular form in (30c) in Buriat comes from the lack of general number readings of unmarked nouns in certain syntactic contexts. In particular, non-plural-marked nouns that appear with certain modifiers (e.g. adjectives) cannot receive number neutral readings; instead they are interpreted as atomic. Similarly, non-plural-marked objects with overt accusative case marking must receive atomic readings. Bylinina and Podobryaev assume that these syntactic constructions require the presence of a NumP, which, in the absence of plural marking, is atomic (30c). On this view, English and Buriat differ only in that DPs in English must project a NumP; there is no number neutral bare form in English.

While this account preserves an implicature-based analysis of exclusive plural for Buriat, it is less appealing for Tiwa, as there is no comparable evidence for a null singular form. In stark contrast to Buriat, number neutral readings of non-plural-marked nouns are available to Tiwa DPs, regardless of their syntactic complexity. This includes exactly the cases where atomic readings are forced in Buriat, namely DPs that are modified by adjectives and those that have overt accusative case marking.¹⁰ One such DP is found in (16) above, where *káshong* ‘dress’ is modified by a nominalized relative clause and an adjective, but still receives a number-neutral reading. The sentence in (31) provides another example. Here, speaker A uses the non-plural marked *khumrái*, modified by the nominalized verb *sípa* ‘small’ and a possessor *Parsinge* ‘Parsing’s’. The entire DP is also marked with accusative case.

- (31) A: Sonrai [Parsing-e síp-a khumrái-gô] pre-ga.
 Sonrai Parsing-GEN small-NMLZ pumpkin-ACC buy-PFV
 ‘Sonrai bought Parsing’s small pumpkin(s).’

to some actual expression in the language. Buccola et al. (2022) have recently argued that alternatives instead need only correspond to a primitive concept, even if the language does not provide a way of encoding that meaning directly. This is an idea worth exploring further for the Tiwa data presented here, but I leave that for future work.

¹⁰Note that Tiwa does, like Buriat, exhibit differential object marking: objects can appear without accusative case (Dawson 2020). Crucially though, unlike in Buriat, the accusative marker does not force atomic readings, as is clearly evident from examples like (2).

B: Khui-tha pre-ga?
 how.many-CL buy-PFV
 ‘How many did she buy?’ [MM 2023.1.23]

In Buriat, such a DP would necessarily receive an atomic reading. In Tiwa, it remains number-neutral. This is clear from the felicity of speaker B’s question asking about the quantity of pumpkins. The structural complexity of this DP does not force an atomic reading, suggesting that such DPs do not require a null SG feature in their NumP in the absence of an overt PL feature.

Of course, it is still possible that Tiwa has a null singular form, as in (30), but that unlike in Buriat, the NumP projection is never required. That is, non-plural-marked nouns are *always* ambiguous between a DP that lacks a NumP projection entirely and one that has a NumP projection with a null SG feature. This, however, could lead to a learnability problem, since there are no contexts that require a null SG feature to be posited. Instead, I assume that bare nouns in Tiwa uniformly lack a SG feature that imposes an atomicity requirement. Instead, bare nouns are always truly number neutral. If this is the right approach, a quantity implicature account as sketched out above cannot account for exclusive readings of plural in Tiwa. This in turn implies that exclusive readings must be derived via a different mechanism, at least in some languages.

4.3 Towards an analysis

In the previous section, I argued against a quantity implicature approach to exclusive readings of plural in Tiwa. While further work is required to detail a full semantic account of Tiwa’s plural marker, in this section I discuss some key considerations.

In contrast to the quantity implicature approach ruled out above, a polysemy approach to plural meaning as outlined in Farkas and de Swart (2010) does not face the same problems for Tiwa. On this approach, plural-marked nouns have two available readings: an inclusive reading and an exclusive reading. A pragmatic principle that favors the strongest meaning in a given context ensures that exclusive readings are generally favored in upward-entailing environments, while inclusive readings are generally favored in downward-entailing environments, unless the context suggests otherwise. Thus the plural-marked nouns in §2 receive exclusive readings, while those in §3.2 receive inclusive readings: the strongest meaning is chosen.

A key fact that any analysis of the Tiwa data must account for is that while plural-marked nouns have inclusive readings, these are not equivalent in meaning to the truly number-neutral readings of bare nouns (cf. Bylinina and Podobryaev’s analysis of Buriat bare nouns and plural-marked nouns in (30)). Crucially, inclusive plurals are not interchangeable with bare nouns if the context suggests there are no possible plural witnesses. For example, while the plural and non-plural forms of the noun are judged by speakers to be equivalent in meaning in a sentence like (32), the plural-marked form of the noun *sí* ‘wife’ in (33a) is judged infelicitous, since its not possible for Samsing to have more than one wife.

(32) Ái korkhyá(-râw) cha.
 1 SG child-PL exist.NEG
 ‘I don’t have children.’ [BM 2023.1.12]

- (33) a. # Samsing-e sí-râw cha.
 Samsing-GEN wife-PL exist.NEG
 ‘Samsing doesn’t have wives.’ [BM 2023.1.13]
- b. Samsing-e sí cha.
 Samsing-GEN wife exist.NEG
 ‘Samsing doesn’t have a wife.’ [BM 2023.1.13]

The use of a plural-marked form in Tiwa, in contrast to the number-neutral bare form, is only licensed when there are possible plural witnesses. Farkas and de Swart (2010) point out similar facts for English and Hungarian plurals, and their polysemy analysis seeks to account for this fact.

Finally, while I have argued that a quantity implicature account of exclusive readings of plurals is untenable for Tiwa, I do not mean to imply that the bare form of the noun and the plural-marked form do not compete pragmatically at all. Plural marking is generally preferred by speakers on human nouns in upward-entailing contexts in the absence of a plurality-entailing numeral or quantifier. Similarly, while inclusive readings of plural nouns in Tiwa emerge easily in questions and under negation, the data are more complicated for plurals in conditional antecedents. In contrast to questions and negative contexts, speakers prefer unmarked nouns to plural-marked nouns for an inclusive reading. The two forms of the noun thus clearly compete pragmatically on some level. Further work is needed to explore the extent of this competition and to offer an analysis.

5 Conclusion

In this paper, I have shown that Tiwa is a general number language with an inclusive plural marker. Further in-depth cross-linguistic work is needed to explore whether any general number language truly has purely exclusive plurality. I additionally have argued that a traditional quantity implicature account of exclusive readings of plurals is difficult to maintain for the Tiwa data, while a polysemy approach along the lines of Farkas and de Swart (2010) is more promising.

References

- Bale, Alan, Michaël Gagnon, and Hrayr Khanjian. 2011. Cross-linguistic representations of numerals and number marking. *Proceedings of Semantic and Linguistic Theory* 20:582–598.
- Buccola, Brian, Manual Križ, and Emmanuel Chemla. 2022. Conceptual alternatives. *Linguistics and Philosophy* 45:265–291.
- Bylinina, Lisa, and Alexander Podobryaev. 2020. Plurality in Buriat and structurally constrained alternatives. *Journal of Semantics* 37:117–128.
- Chierchia, Gennaro. 1998. Reference to kinds across languages. *Natural Language Semantics* 6:339–405.
- Corbett, Greville. 2000. *Number*. Cambridge University Press.
- Dawson, Virginia. 2018. Countability and plurality in a classifier language. Paper presented at the 92nd meeting of the Linguistic Society of America, Salt Lake City.
- Dawson, Virginia. 2020. Existential quantification in Tiwa: Disjunction and indefinites. Doctoral Dissertation, University of California, Berkeley.

- Dawson, Virginia, and Peter Jenks. 2023. Describing definites and indefinites. In *Key topics in language documentation and description*, ed. Peter Jenks and Lev Michael, Language Documentation and Conservation Special Publication No. 26, 99–174. Honolulu: University of Hawai'i Press.
- Eberhard, David M., Gary F. Simons, and Charles D. Fennig, ed. 2023. *Ethnologue: Languages of the world, twenty-sixth edition*. Dallas, Texas: SIL International. Online edition: <http://www.ethnologue.com>.
- Farkas, Donka, and Henriëtte de Swart. 2010. The semantics and pragmatics of plurals. *Semantics and Pragmatics* 3:1–54.
- Fox, Danny, and Roni Katzir. 2011. On the characterization of alternatives. *Natural Language Semantics* 19:87–107.
- Joseph, U.V., ed. 2014. *Tiwa-English dictionary with English-Tiwa index*. Shillong: Don Bosco Centre for Indigenous Cultures. Associate editors: Horsing Kholar, Juliana Maslai, Alfred Maslai, Bibiana Maslai, and Simon Mithi.
- Kang, Beom-Mo. 1994. Plurality and other semantic aspects of common nouns in Korean. *Journal of East Asian Linguistics* 3:1–24.
- Katzir, Roni. 2007. Structurally-defined alternatives. *Linguistics and Philosophy* 30:669–690.
- Li, Yen-Hui Audrey. 1999. Plurality in a classifier language. *Journal of East Asian Linguistics* 8:75–99.
- Martí, Luisa. 2020. Inclusive plurals and the theory of number. *Linguistic Inquiry* 51:37–74.
- Matthewson, Lisa. 2004. On the methodology of semantic fieldwork. *International Journal of American Linguistics* 70:369–415.
- Nakanishi, Kimiko, and Satoshi Tomioka. 2004. Japanese plurals are exceptional. *Journal of East Asian Linguistics* 13:113–140.
- Post, Mark W., and Robbins Burling. 2017. The Tibeto-Burman languages of Northeast India. In *The Sino-Tibetan languages*, ed. Graham Thurgood and Randy J. LaPolla, 213–242. Routledge, 2nd edition.
- Renans, Agata, Yağmur Sağ, F. Nihan Ketrez, Lyn Tieu, George Tsoulas, Raffaella Folli, Hana de Vries, and Jacopo Romoli. 2020. Plurality and crosslinguistic variation: an experimental investigation of the Turkish plural. *Natural Language Semantics* 28:307–342.
- Rullmann, Hotze, and Aili You. 2006. General number and the semantics and pragmatics of indefinite bare nouns in Mandarin Chinese. In *Where semantics meets pragmatics*, ed. Klaus von Heusinger and Ken Turner, 175–196. Brill.
- Sauerland, Uli, Jan Andersen, and Kazuko Yatsushiro. 2005. The plural is semantically unmarked. In *Linguistic evidence: Empirical, theoretical, and computational perspectives*, ed. Stephan Kepser and Marga Reis, 413–434. Berlin: Mouton de Gruyter.
- Spector, Benjamin. 2007. Aspects of the pragmatics of plural morphology: on higher-order implicatures. In *Presuppositions and implicatures in compositional semantics*, ed. Uli Sauerland and Penka Stateva, 243–281. New York: Palgrave-Macmillan.
- Zweig, Eytan. 2009. Number-neutral bare plurals and the multiplicity implicature. *Linguistics and Philosophy* 32:353–407.

Kumyk Verb Classification: Event and Argument Structure¹

Anastasija Gruzdeva — Institute of Linguistics, RAS / Lomonosov MSU

Anna Alhazova — Lomonosov MSU

Anna Golovnina — Lomonosov MSU

Regina Nasyrova — Lomonosov MSU

Feudor Sadkovsky — Institute of Linguistics, RAS / Lomonosov MSU

Abstract. This paper provides a classification of Terek Kumyk verbs in terms of event and argument structure. The classes are built on three main parameters: stativity / dynamicity, degree of agentivity, and presence and type of a result state in the event structure of a verb. The results of the classification can be further used for making various generalizations about syntax and semantics of Terek Kumyk.

1 Introduction

This paper aims at identifying verb classes in Kumyk, which belongs to the Kipchak sub-branch of the Turkic languages, along with Karachay-Balkar, Karaim, and Crimean Tatar (Gadžiahmedov et al. 2014). The study is based on a database of 61 verbs collected by the authors during the expeditions to the village of Predgornoje in the Republic of North Ossetia–Alania in August 2022 and 2023. The data was elicited from around ten Terek Kumyk speakers aged 15–65, most of them are female teachers of the local school. Fig. 1 shows an example of the entry interface for the verb *uxlamaq* ‘to sleep’ in our database.

ID	2
Verb	ухламакъ
Translation	спать
Diathesis	NOM
IsDerivate	
Assignee	Настя Г.
BaseExample	1. Заур ухлады 2. Заур ухлай
BaseTranslation	1. Заур заснул. 2. Заур спит.
Consultants	Рукуят, Карина, Зарема
Comments	

Actionality	Transitivity	Agentivity	ResState
Actionality	ES S P, S P		
PerfTelicExample	1. Заур эки сагъатдан ухлады		
PerfTelicTranslation	1. Заур заснул через 2 часа (после чего-то) / %за		
PerfAtelicExample	Заур эки сагъат ухлады		
PerfAtelicTranslation	Заур спал/проспал/засыпал 2 часа.		
ImperfTelicExample	Заур эки сагъатдан ухлай		
ImperfTelicTranslation	Заур засыпает за 2 часа (по жизни, не прямо)		
ImperfAtelicExample	Заур эки сагъат ухлай		
ImperfAtelicTranslation	Заур спит 2 часа (сейчас / вообще).		
Consultants	Рукуят, Карина, Зияудин, Зоя, Оксана		
Comments			

Figure 1: Entry example

¹The study has been supported by RSF, project #22-18-00285 “Scalarity in the Grammar and Lexicon: a semantic and typological study” at the Lomonosov Moscow State University.

The main objective of the study was to classify a sample of 61 verbs in terms of argument and event structure. The parameters of the classification are discussed in Section 2. The resulting classification of the verb sample for Terek Kumyk is presented in Section 3. Possible practical applications and future prospects of the study are given in Section 4.

2 Classification parameters

Three semantic parameters have been taken into account: stativity/dynamicity, degree of agentivity, and existence and type of a resulting state (Fig. 2). Our classification follows Privoznov and Gruzdeva (2018) where these three parameters are used to establish verb classes in Barguzin Buryat. In addition, each verb was assigned one of the eventuality types (see Section 2.1) and a diathesis.

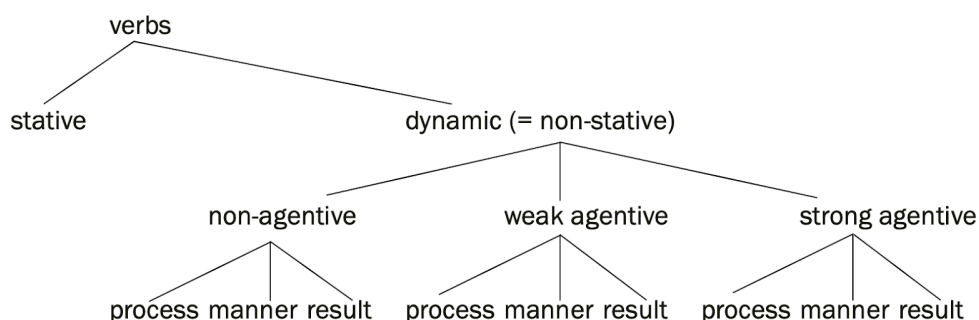


Figure 2: Verb classes by Privoznov and Gruzdeva (2018)

As Fig. 2 shows, all verbs fall into two categories: stative verbs and dynamic verbs (also known as non-stative, trivially). The next subdivision has to do with the semantic property of agentivity: verbs that have it are called agentive and vary in degree of agentivity (there are weak agentive verbs and strong agentive ones); verbs lacking agentivity are called non-agentive. At the same time, all dynamic verbs can be divided into two groups: verbs with a resulting state and without it. Depending on the type of the resulting state, the latter group is further subdivided into manner verbs and result verbs, whereas verbs lacking a resulting state are called processes.

Each of the parameters is discussed in more detail in the following subsections.

2.1 Eventuality classification

The eventuality classification reflects differences between verbs in terms of aspectual potential and available interpretations. The inventory of eventuality types in our work is based on Tatevosov (2016) and takes into account the behavior of a predicate in imperfective and perfective contexts. An imperfective context is provided by the verb forms with the imperfective affix *-A/-j*; for a perfective context the so-called ‘categorical past’ (Gadžiahmedov et al. 2014) marker *-dI* has been chosen. For each of these two main aspectual forms of every verb in the sample, available actional

meanings were identified. The study assumes six possible meanings: atelic meanings S (state), P (process), and MP (multiplicative process), as well as telic meanings ES (entry into a state), EP (entry into a process), EMP (entry into a multiplicative process). We used adjuncts like *eki sahat* ‘for two hours’ and *eki sahat-dan* ‘in two hours’ to set atelic and telic contexts, respectively, following Dowty (1979). The adjunct *eki-de* ‘at two (o’clock)’ was also used to elicit entry into a process (EP) meanings for some verbs.

For instance, the eventuality type of the verb *uxlamaq* ‘to sleep’ is <ES P S, P S>—this notation represents possible meanings of the verb in perfective and imperfective contexts separated by a comma. In the perfective context an entry into a state interpretation ‘attain a state of sleeping’ (2), along with a state ‘be in a state of sleeping’ (1.1) and a process ‘be in a process of falling asleep’ (1.2) interpretations are attested, whereas only the latter two are available in the imperfective context, see examples (3.1) for S and (3.2) for P. Expectedly, actual-durative, or episodic, telic interpretations are not available in the imperfective context (4). Habitual readings, like in (4.1), are not relevant for our test because they simply describe individual states, i. e. permanent properties (see Lyutikova et al. (2006)) and episodic aspectual characteristics do not apply to them. The “scheduled future” meaning (4.2) does not exactly refer to actual-durative situations either.

perfective context: S, P

- (1) *Zaur eki sahat uxla-di*
 Zaur.NOM two hour.ACC sleep-PST
 1. ‘Zaur slept for 2 hours.’
 2. ‘Zaur was falling asleep for 2 hours.’

perfective context: ES

- (2) *Zaur eki sahat-dan uxla-di*
 Zaur.NOM two hour-ABL sleep-PST
 1. ‘Zaur fell asleep in 2 hours.’
 2. ‘Zaur fell asleep 2 hours later.’

imperfective context: S, P

- (3) *Zaur eki sahat uxla-j*
 Zaur.NOM two hour.ACC sleep-IPFV
 1. ‘Zaur has (already) been sleeping for 2 hours.’
 2. ‘Zaur has (already) been falling asleep for 2 hours.’

imperfective context: ∅

- (4) *Zaur eki sahat-dan uxla-j*
 Zaur.NOM two hour-ABL sleep-IPFV
 1. ‘Zaur (usually) falls asleep in 2 hours.’
 2. ‘Zaur is falling asleep in 2 hours (according to the schedule).’

2.2 Stativity vs. dynamicity

The stativity/dynamicity parameter relies on the eventuality type of verbs. If the eventuality type of a verb contains an ‘S’, we consider such a verb stative (for subtypes of stative verbs, see Tatevosov (2016)).

The distinction between states and processes is based on multiple criteria, as described in Lyutikova et al. (2006), Tatevosov (2016), etc. Inherently, S-meaning corresponds to a situation that does not change over time and does not require constant energy flow. Moreover, propositions with stative predicates retain their truth value within any subinterval of the situation duration, however small it may be, and even at any point of the state duration. Formally, when X is a situation and X_0 is a subinterval of X of length t , if a proposition P , describing X , also holds for X_0 with any t -value (including 0), we can posit that X is a state. Furthermore, due to this definition states can never have Agents as their subjects; therefore, the two other parameters of the classification, the degree of agentivity as well as existence and type of a resulting state, are only applicable to non-stative verbs.

The eventuality type of the verb *uxlamaq* ‘to sleep’ (<ES P S, P S>) contains ‘S’ in both contexts, so it is considered a stative verb in our classification. Indeed, the situation of sleeping is homogenous and does not require constant energy influx. Any subinterval of sleeping can be truthfully described with the verb ‘to sleep’; moreover, the proposition ‘Zaur was sleeping’ is true at any point of the sleeping state duration.

Stative verbs are usually non-agentive: they describe a situation that cannot be controlled by the subject, e.g. *to see*. However, some stative verbs are underspecified in terms of agentivity: For such verbs, the situation described by the predicate can optionally be controlled by the subject but not necessarily, e.g. *to sit* (Tatevosov 2016). In any case, the agentivity parameter of our classification is not applicable for statives. Subjects of stative verbs are considered Holders of the State—this term refers to Experiencers and arguments of one-place predicates such as *to sit* (see Ramchand (2006)).

Non-stative verbs are called dynamic (Lyutikova et al. 2006). For example, *ačmaq* ‘to open’ has no S in its eventuality type <ES P, P>, which makes it dynamic. In the perfective context, two meanings are possible: entry into a state of being opened (of an object, namely the door in (6) and (7.1)) and the process of opening something (5), while the meaning of entry into a process is not available, as (7.2) shows. In the imperfective context, we can only get a process interpretation (8).
perfective context: P

- (5) *Mustafa eşik-ni eki sahat aç-di*
Mustafa.NOM door-ACC two hour open-PST
‘Mustafa was opening the door for 2 hours.’

perfective context: ES

- (6) *Mustafa eşik-ni eki sahat-dan aç-di*
Mustafa.NOM door-ACC two hour-ABL open-PST
1. ‘Mustafa opened the door 2 hours later.’
2. ‘Mustafa opened the door in 2 hours.’

perfective context: ES, *EP

- (7) *Mustafa eki-de ešik-ni ač-di*
 Mustafa.NOM two-LOC door-ACC open-PST
 1. ‘Mustafa opened the door at 2 (o’clock).’
 2. *‘Mustafa started opening the door at 2 (o’clock).’

imperfective context: P

- (8) *Mustafa ešik-ni eki sahat ač-a*
 Mustafa.NOM door-ACC two hour open-IPFV
 ‘Mustafa has (already) been opening the door for 2 hours.’

Some dynamic verbs can take Agents as their subjects, this provides the basis for the next classification parameter. Regardless of the agentivity type, dynamic verbs are further classified by presence and type of a resulting state in their event structure.

2.3 Agentivity

Agentive predicates denote situations that are caused by actions of the subject and are under its immediate control. Agentive verbs are usually said to have a causing subevent in their syntax-semantics interface, e. g. heads like DO in Dowty (1979), Voice in Kratzer (1996), etc. We have used two types of criteria to evaluate the agentivity of a verb. Firstly, only agentive verbs are compatible with agent-oriented modifiers, such as phrases expressing purpose (9), modifiers that mean ‘unintentionally’ (10) and ‘in a hurry’ (11).

(9) shows the difference between an agentive verb *ačmaq* ‘to open’ (9a) and a non-agentive verb *xurillamaq* ‘to snore’ (9b). The former can be modified with a purpose expression *jel qaqtava* ‘(in order) to ventilate’. On the contrary, a purpose expression *qawırba-si-na burulmaq sajalı* ‘(in order) to roll on his side’ renders the sentence with the non-agentive verb infelicitous.

- (9) a. *Mustafa ešik-ni ač-di jel qaqtava*
 Mustafa.NOM door-ACC open-PST wind.ACC knock-INF
 ‘Mustafa opened the door to ventilate.’
 b. **Murat xurilla-di qawırba-si-na burulmaq sajalı*
 Murat.NOM snore-PST side-3-DAT turn-INF for
 Int.: ‘Murat snored to roll on his side.’

As (10a) and (10b) demonstrate, agentive verbs like *ačmaq* ‘to open’ and *ašamaq* ‘to eat’ are compatible with adverbials *bilmej* and *xaparlamıj*², which both mean ‘unintentionally, accidentally, unknowingly’. (10c) illustrates incompatibility of a non-agentive verb *ölmek* ‘to die’ with such a

²The inner form of this word remains unclear, although it is most likely a negative imperfective converb like *bilmej* derived from the verb *xaparlamaq*. However, this alleged word is not registered in Kumyk dictionaries (e.g. Agamov (2018)), they only contain *xabarlamaq* ‘to converse, talk’.

modifier. The non-agentive verb *tammaq* ‘to drip’ is incompatible with *xaparlamij* as well (10d). However, some speakers considered examples such as (10d) felicitous (in this instance—with *bilmej* instead of *xaparlamij*). In such cases unintentionality adverbials are controlled by an Agent or a Causer, assumed in the context outside of the sentence.

- (10) a. *Mustafa bil-me-j ešik-ni ač-di*
 Mustafa.NOM know-NEG-IPFV door-ACC open-PST
 ‘Mustafa accidentally, unintentionally opened the door. (It was not properly closed, he leaned on it.)’
- b. *Umar bil-me-j/ xaparlamij žibin(-ni) aša-di*
 Umar.NOM know-NEG-IPFV unintentionally fly-ACC eat-PST
 ‘Umar accidentally ate a fly.’
- c. **Salix bil-me-j öl-di*
 Salikh know-NEG-IPFV die-PST
 Int.: ‘Salikh died unintentionally.’
- d. *suw bil-me-j/ *xaparlamij kurške-ge tam-ni*
 water.NOM know-NEG-IPFV unintentionally cup-DAT drip-PST
 ‘Water accidentally/*without thinking dripped into the cup.’

In (11a) the agentive verb *ačmaq* ‘to open’ forms an acceptable sentence with the adverbial *albasap* ‘in a hurry’, unlike the non-agentive verb *xurillamaq* ‘to snore’ (11b):

- (11) a. *Mustafa albasap ešik-ni ač-di*
 Mustafa.NOM hurry-CONV door-ACC open-PST
 ‘Mustafa opened the door in a hurry.’
- b. **Zaur albasap xurilla-di*
 Zaur.NOM hurry-CONV snore-PST
 Int.: ‘Zaur snored in a hurry.’

The second criterion for evaluating agentivity that we used is the ability to form imperatives. Usually, only imperatives from agentive verbs can have the semantics of an order or request (12). For non-agentive verbs, it is either impossible to derive an imperative or its meaning is different (e.g. optative). For instance, the imperative form of the agentive verb *ačmaq* ‘to open’ (12) is normal and can be used in the direct imperative meaning whereas the imperative form of the non-agentive *ölmek* ‘to die’ is either judged infelicitous or interpreted as optative.

- (12) a. *Ač!*
 open.IMP
 ‘Open!’
- b. **Öl!*
 die.IMP
 Int.: ‘Die!’ (as a command)

If most tests indicated that a verb is agentive³, we further checked whether the verb allows an Effector as its subject or only an Agent. In the first case, the verb is considered to be weak agentive (cf. verbs with optional DO in Levin and Rappaport Hovav (1994), in the second case it is called strong agentive. For example, in (13) the weak agentive verb *ačmaq* ‘to open’ can have either an Agent *Mustafa* (13a) or an inanimate Effector *jel* ‘wind’ (13b) as its subject whereas the strong agentive *almaq* ‘take (sth off)’ only allows for the former (14).

- (13) a. *Mustafa ešik-ni ač-dî*
Mustafa.NOM door-ACC open-PST
‘Mustafa opened the door.’
- b. *Jel ešik-ni ač-dî*
wind.NOM door-ACC open-PST
‘The wind opened the door.’
- (14) a. *Mustafa kiniške-ni polka-dan al-dî*
Mustafa.NOM book-ACC shelf-ABL take.off-PST
‘Mustafa took the book off the shelf.’
- b. **jel kiniške-ni al-dî*
wind.NOM book-ACC take.off-PST
Int.: ‘The wind took the book off.’

2.4 Result state: presence and type

The last parameter of our classification is the presence and type of a result state. Verbs with a result state allow two interpretations in sentences with negation or decomposition adverbs such as *jañidan* ‘again’ or *(bir) dabi da* ‘once more’ and *poçti* ‘almost’. In negative sentences either the whole proposition or the resulting state can be under negation (Lyutikova et al. (2006), Privoznov and Gruzdeva (2018)). ‘Again’-modifiers are known for their ability to bring about the so-called repetitive/restitutive ambiguity, see, for example, von Stechow (1996), Bale (2007), Xu (2016). Under the repetitive interpretation, an entire eventuality described by the predicate is in the scope of the modifier, while in the restitutive interpretation, it is only the result that is in the scope of ‘again’. Sentences that contain verbs with no result state (denoting processes) only have the repetitive interpretation. When a predicate with a result state is modified with ‘almost’, again, two readings are possible: counterfactual—when the whole predicate is in the scope of the modifier—and scalar—when only the result state is canceled (Dowty (1979), Xu (2016)).

(15) shows that both kinds of interpretations are available for the verb *ačmaq* ‘to open’. In (15a) negation can be interpreted either below the external argument (15a.2) or above it (15a.1).

³Verbs denoting actions typically performed by animals are a special case: They do not pass most of the agentivity tests due to people’s perception of animals and their consciousness. Anyhow, animals are animate and therefore Agents, not Effectors, which is why such verbs were considered agentive nevertheless.

In the former case only the resulting state is negated, in the latter case the whole proposition is in the scope of negation. In (15b) the first interpretation is repetitive: The whole situation repeats itself with no change of participants; the second interpretation is restitutive: The result state of the door being opened repeats with help of another Agent. (15c) demonstrates counterfactual/scalar ambiguity of ‘almost’: In the first reading nothing happened, the Agent has not initiated the action yet though he got close to it; in the second reading the Agent started performing the action but the resulting state of the door has not been achieved.

- (15) a. *Mustafa ešik-ni ač-ma-dî*
Mustafa.NOM door.ACC open-NEG-PST
1. ‘Mustafa didn’t open the door. (He did not even touch it.)’
2. ‘Mustafa didn’t open the door. (He tried to open it but was not able to.)’
- b. *Mustafa jaŋıdan/ bir da vı ešik-ni ač-dî*
Mustafa.NOM again one again ADD door.ACC open-PST
1. repetitive
‘(Mustafa had already opened the door and then) Mustafa opened the door again.’
2. restitutive
‘(Kerim had already opened the door and after that) Mustafa opened the door again.’
- c. *Mustafa ešik-ni poçti ač-dî*
Mustafa.NOM door.ACC almost open-PST
1. counterfactual
‘Mustafa almost opened the door: He got close to start opening the door but has not started yet.’
2. scalar
‘Mustafa almost opened the door: He started opening it but the door is not fully opened yet.’

We call verbs with no result state processes. *Xurillamaq* ‘to snore’ is an example of such a verb: negation (16a) and ‘again’-modifiers (16b) do not bring about ambiguity in sentences containing it.

- (16) a. *Murat xurilla-ma-dî*
Murat.NOM snore-NEG-PST
1. ‘Murat did not snore.’
2. *‘Murat snored but didn’t snore himself out.’
- b. *Murat jaŋıdan/ da vı xurilla-dî*
Murat.NOM again again ADD snore-PST
1. ‘Murat snored again. (He had snored and stopped before.)’
2. *‘(Kerim had snored before, then) Murat snored again.’

The ‘again’ test is not applicable to intransitive verbs, so we used adverbials of temporary state on them instead: modifiers like *eki sahat-va* ‘for two hours’ can denote temporariness of state (and not

duration of action, at least not only) when modifying verbs with a result state (Dowty 1979). For example, in (17) it is Kerim’s state of staying in a certain location that is modified by the temporal expression, not the process of coming.

- (17) *Kerim eki sahat-va gel-di*
 Kerim.NOM two hour-DAT come-PST
 ‘Kerim came (to stay) for two hours.’

Among the verbs with a result state, following Rappaport Hovav and Levin (1998), we distinguish between manner and result verbs. Manner verbs usually describe the way in which the action was performed but rarely entail the result that was achieved with this action. Result verbs, on the contrary, often specify a particular result but not the way the action was carried out. We used the ability of a predicate to omit a direct object as a test. Only manner verbs allow to do so without bringing about ungrammaticality as their “result state is believed to build up with a direct object and is considered to be an optional part of the verb semantics” (Privoznov and Gruzdeva 2018). For example, *ačmaq* ‘open’ is a result verb and cannot be used without its direct object, as we can see from (18b). *Ašamaq* ‘eat’ is an example of a manner verb. In (18b), we can see that its direct object can be omitted without turning the sentence ungrammatical.

- (18) a. *Mustafa ešik-ni aç-di*
 Mustafa.NOM door-ACC open-PST
 ‘Mustafa opened the door.’
 b. *Mustafa aç-di*
 Mustafa.NOM open-PST
 #‘Mustafa opened.’ (incomplete sentence)
- (19) a. *Umar šorpa aša-di*
 Umar.NOM soup.ACC eat-PST
 ‘Umar ate the soup.’
 b. *Umar aša-di*
 Umar.NOM eat-PST
 ‘Umar ate.’ (correct sentence)

However, this criterion is not applicable to intransitive verbs since they do not have a direct object in their diathesis. Another test that is sometimes used to distinguish between manner and result verbs is the ability to be modified by secondary predicates, such as adjectives and oblique objects (e.g. *Terry swept the floor **clean***, Rappaport Hovav and Levin (1998)). Verbs that allow secondary predicates in their structure are considered to be manner verbs. Unfortunately, we have not been able to use this test on Kumyk because this language does not seem to allow resultative secondary predication at all. For this reason, we are yet to find an appropriate test to distinguish between intransitive result and manner verbs.

3 Terek Kumyk verbs

The results of the classification are given below. Some of the characteristics are italicized: it means that for now the data on this aspect is either controversial or scarce.

	Verb	Translation	Diathesis	Eventuality	Class
1	bilmek	to know	NOM ACC	ES, S	stative
2	uxlamaq	to sleep	NOM	ES P S, P S	stative
3	bojamaq	to paint	NOM ACC	ES P, P	weak agentive result verb
4	aşmaq	to eat	NOM ACC	ES P, P	weak agentive manner verb
5	žuwmaq	to wash	NOM ACC	ES P, P	weak agentive result verb
6	jazmaq	to write	NOM ACC	ES P, P	strong agentive manner verb
7	xurillamaq	to snore	NOM	EP P, P	non-agentive process verb
8	uçmaq	to fly	NOM	EP P, P	weak agentive process verb
9	qiçirmaq_1	to shout	NOM	EP ES P, P	strong agentive process verb
10	açmaq	to open	NOM ACC	ES P, P	weak agentive result verb
11	kesmek	to cut	NOM ACC	ES P, P	weak agentive result verb
12	tazalamaq	to clean	NOM ACC	ES P, P	weak agentive result verb
13	suwumaq	to cool down	NOM	ES P, P	non-agentive verb with a result state
14	açımaq	to turn sour	NOM	ES <i>P, P</i>	non-agentive verb with a result state
15	turmaq	to stand	NOM	ES S, S	stative
16	işlemek	to work	NOM	EP P, P	weak agentive process verb
17	barmaq	to go	NOM (DAT)	EP ES P, P	strong agentive verb with a result state
18	jaşamaq	to live	NOM (LOC)	ES S, S	stative
19	atmaq	to throw	NOM ACC (DAT)	ES MP P, MP P	weak agentive result verb
20	aşarmaq	to turn white, pale	NOM	ES S, S	non-agentive verb with a result state
21	görmek	to see	NOM ACC	ES S, S	stative
22	qaramaq	to look, watch	NOM DAT	ES <i>P, P</i>	strong agentive result verb
23	olturmaq	to sit	NOM (LOC)	ES S, S	stative
24	izlemek	to search	NOM ACC	<i>EP P, P</i>	strong agentive process verb
25	çapmaq	to run	NOM	EP P, P	strong agentive process verb
26	tapmaq	to find	NOM ACC	ES, -	strong agentive result verb
27	uwmaq	to crumble	NOM ACC	ES P, P	strong agentive manner verb
28	jibermek	to send	NOM ACC DAT	EP ES P, P	strong agentive result verb
29	külemek	to laugh	NOM (DAT)	EP P, P	strong agentive process verb
30	gelmek	to come	NOM (DAT)	ES, -	weak agentive verb with a result state
31	qaqmaq_1	to knock	NOM ACC/ DAT	EP ES P, P	weak agentive process verb
32	ojnamaq	to play	NOM ACC	EP ES P, P	strong agentive manner verb

33	oxumaq	to read	NOM ACC	ES P, P	strong agentive manner verb
34	süjmək	to love	NOM ACC	ES S, S	stative
35	eşitmək	to hear	NOM ACC	ES S, S	stative
36	tişləməq	to listen	NOM DAT	ES P, P	strong agentive process verb
37	bermək	to give	NOM ACC DAT	ES P, P	strong agentive result verb
38	ulumaq	to howl	NOM	EP ES P, P	strong agentive process verb
39	quwalamaq	to chase (off)	NOM ACC	EP ES P, P	weak agentive <i>manner</i> verb
40	tammaq	to drip	NOM	EMP ES MP, MP	non-agentive process verb
41	haplamaq	to bark	NOM	ES EMP MP, MP	strong agentive process verb
42	joluqmaq_1	to meet	NOM COM/ NOMpl	ES P, P	weak agentive process verb
43	urmaq_1	to throw	NOM ACC (DAT)	ES MP P, MP P	<i>weak</i> agentive result verb
44	jüzmək	to swim	NOM	EP ES P, P	weak agentive verb <i>with a result state</i>
45	öpmeq	to kiss	NOM ACC	ES MP P, MP P	strong agentive <i>result</i> verb
46	ajtmaq	to speak	NOM ACC DAT	EP ES P, P	strong agentive <i>manner</i> verb
47	čeçmək	to untie	NOM ACC	EP ES P, P	<i>strong</i> agentive result verb
48	ölmək	to die	NOM	ES P, P	non-agentive verb with a result state
49	buzmaq	to spoil	NOM ACC	EP ES P, P	weak agentive <i>result</i> verb
50	unutmaq	to forget	NOM ACC	ES P, P	<i>non-agentive result</i> verb
51	gijmək	to put on	NOM ACC	ES P, P	strong agentive result verb
52	tartmaq	to pull	NOM ACC	EP ES P, P	strong agentive result verb
53	almaq	to take off	NOM ACC	ES, P	weak agentive result verb
54	jötkürmək	to cough	NOM	ES EMP MP, MP	strong agentive process verb
55	toxtamaq	to wait	NOM (DAT)	S, S	stative
56	xapmaq	to bite	NOM ACC	EMP ES MP P, MP P	weak agentive result verb
57	qurmaq	to build	NOM ACC	ES P, P	strong agentive <i>manner</i> verb
58	qaqmaq_2	to hammer in	NOM ACC	ES P, P	weak agentive result verb
59	joluqmaq_2	to meet; to agree	NOM DAT	ES, -	non-agentive result verb
60	urmaq_2	to hit	NOM DAT	EMP ES MP P, MP P	weak agentive process verb
61	qiçirmaq_2	to call, shout out to	NOM DAT	ES P, P	strong agentive process verb

4 Conclusion and prospects

This paper presents the results of classification of Kumyk verbs in terms of argument and event structure. Three main parameters were taken into account: stativity / dynamicity, degree of agentivity, and presence and type of a result state. Apart from that, we investigated aspectual characteristics of the verbs in the sample and assigned each one of them an eventuality type.

Whereas some verbs are more obvious to classify (e.g., ‘to sleep’ is unlikely to be anything but stative in any language), other verbs can vary across languages in terms of their event structure: for example, ‘to paint’ is a result verb in English but a manner verb in Buryat. For this reason, it is always good to have a reliable language-specific verb classification to base generalizations upon.

Class membership can predict a number of aspects of the grammatical behavior of a verb, including its ability to undergo various argument structure changing operations, aspectual and event structural modifications. For example, as shown in Lyutikova et al. (2006), in Balkar (a Turkic language genetically very close to Kumyk), non-agentive processes can only form contact causatives, while agentive processes and result verbs can only form distant causatives. Other examples of such generalizations about valency-changing operations can be found in Privoznov and Gruzdeva (2018): In Buryat, for instance, it is impossible to derive decausatives from strong agentive verbs, and non-agentive verbs can only form causal passives. As far as aspectual modifications are concerned, Lyutikova et al. (2006) show that Balkar constructions with the auxiliary verb *tura-* ‘stand’, normally describing process phases of single situations, behave differently depending on a class of a base verb: for example, in punctual verbs such constructions give rise to an additional dynamic phase, non-existent in simple imperfective forms.

No such generalizations have been made for Kumyk verbs yet but, hopefully, our database and classification will provide a basis for that.

References

- Agamov, Z. (2018). *Kumyk dictionary (1.6)*. [Mobile app, accessed 17.10.2023].
- Bale, A. C. (2007). Quantifiers and verb phrases: An exploration of propositional complexity. *Natural Language & Linguistic Theory*, 25:447–483.
- Dowty, D. R. (1979). *Word meaning and Montague grammar: The semantics of verbs and times in generative semantics and in Montague’s PTQ*. Dordrecht: Reidel.
- Gadžiahmedov, N. E. et al. (2014). *Sovremennyj kumyjskij jazyk [Modern Kumyk]*. Makhachkala: Dagestan Federal Research Center, RAS.
- Kratzer, A. (1996). Severing the external argument from its verb. In Rooryck, J. and Zaring, L., editors, *Phrase Structure and the Lexicon*, pages 109–137. Dordrecht, Kluwer Academic Publishers.
- Levin, B. and Rappaport Hovav, M. (1994). *Unaccusativity: At the syntax-lexical semantics interface*, volume 26. MIT press.
- Lyutikova, E. A., Tatevosov, S. G., et al. (2006). *Struktura sobytija i semantika glagola v karačajevo-balkarskom jazyke [Event structure and verb meaning in Karachay-Balkar]*. Moscow: Gorky Institute of World Literature, RAS.

- Privoznov, D. and Gruzdeva, A. (2018). Aktantnyje derivacii i argumentnaja struktura [valency-changing operations and argument structure]. In Tatevosov, S. G. et al., editors, *Èlementy burâtskogo jazyka v tipologičeskom osvešenii. [Elements of Buryat in a typological perspective]*.
- Ramchand, G. (2006). *Verb meaning and the Lexicon: A First Phase Syntax*, volume 116. University of Tromsø, CASTL.
- Rappaport Hovav, M. and Levin, B. (1998). Building verb meanings. *The projection of arguments: Lexical and compositional factors*, pages 97–134.
- Tatevosov, S. G. (2016). *Glagol'nyje klassy i tipologija akcional'nosti [Verb classes and the typology of actionality]*. Moscow: JaSK.
- von Stechow, A. (1996). The different readings of wieder 'again': A structural account. *Journal of semantics*, 13(2):87–138.
- Xu, T. (2016). Almost again: On the semantics and acquisition of decomposition adverbs.

Variable Modal Strength in Afrikaans and Samoan: Deriving Strong Necessity from Weak Necessity¹

Siena Weingartz & Vera Hohaus— The University of Manchester

Abstract. We discuss a new pattern in the cross-linguistic typology of modal strength, relating to the distinction between weak and strong necessity, based on data from original fieldwork. In both Afrikaans (Indo-European, Germanic; South Africa) and Samoan (Austronesian, Oceanic; American Samoa, Independent State of Samoa), the distinction may be left morphologically unmarked. We suggest that this variability in strength can be explained under a uniform analysis where the relevant expressions are weak necessity modals that allow for the secondary ordering source to be empty, unlike their English counterparts.

1 Introduction

Descriptively, modal expressions can be characterised along three dimensions of meaning: A first dimension of force (relating to the difference between possibility and necessity), a second dimension of flavour (most broadly, relating to the distinction between epistemic and other considerations behind the modal claim), and a third dimension of strength (relating to the perceived strength of the possibility or necessity). Our focus in this paper is on this third and lesser studied dimension of strength, and the perceived difference between weak and strong necessity that the example in (1) from a cartoon by Carolita Johnson for *The New Yorker* picks up on.

- (1) *Employees **must** wash their hands. Non-employees really **ought** to wash their hands, too.*
(see also von Stechow & Iatridou, 2008, p. 115)

The paper seeks to make a contribution to two broad underlying research questions relating to the mapping between form and meaning when it comes to weak necessity: First, how are such strength distinctions encoded across languages? Second, how can the cross-linguistic data inform the semantic analysis of modal strength? Previewing our contribution to the first question, we identify a new pattern in the morpho-syntactic typology of weak necessity, based on data from original fieldwork: In Afrikaans (Indo-European, Germanic) and Samoan (Austronesian, Oceanic), the distinction between weak and strong necessity may systematically be left unmarked, as illustrated in (2) and (3); Afrikaans *moet* and Samoan *tatau* exhibit variable strength.² Previewing our contribution to the second question, we propose to capture this observed variability under an analysis

¹We are indebted to the native speakers involved in this research: *Baie dankie! Fa'afetai!* For valuable feedback and discussion, we would also like to thank Nadine Bade, Martina Faller, Andrew Koontz-Garboden, Manfred Krifka, Viola Schmitt, Ryan Walter Smith, Frank Sode, Malte Zimmermann as well as audiences in Berlin and Potsdam. SW is funded by the Arts and Humanities Research Council (#AH/R012792/1). VH would want to gratefully acknowledge the 2022 fellowship funding from the Leibniz-Centre for General Linguistics (ZAS), Berlin.

²Abbreviations used in glosses: AV = actor voice, CIRC = circumstantial, COMP = complementiser, CND = common noun determiner, COND = conditional, COUNTER = counter to expectation, DEF = definite, DET = determiner, ERG = ergative, FOC = focus, IMPF = imperfective, INCH = inchoative, INDEF = indefinite, INF = infinitive, INFER = epistemic inferential, IRR = irrealis, MID = middle intransitivizer, MOD = modal, NEC = necessity, NEG = negation, NOM = nominaliser, OOC = out of control, PFV = perfective, PL = plural, POS = possibility, POSS = possessive, PREP = preposition, PRN = pronoun, PRT = particle, PST = past, REL = relative, ROOT = root,

within a domain-restriction approach to weak necessity (von Fintel and Iatridou, 2008, 2023; Rubinstein, 2012, 2021; Vander Klok and Hohaus, 2020): The relevant necessity modal expressions are weak necessity modals, but allow for their secondary ordering source to be empty, unlike their English counterparts *should* and *ought*.

- (2) *Werkers moet hande was. Nie-werkers moet ook hulle hande was.* – Afrikaans –
workers NEC hands wash not-workers NEC also their hands wash
'Employees must wash hands. Non-employees should also wash their hands.'
- (3) *Mo tagata fai=galuega: E tatau ona fufulu mamā lima.* – Samoan –
for people make=work TAM NEC that wash.PL clean hand
Mo le mamalu lautele: E matuā tatau foi ona fufulu mamā lima.
for the dignity general TAM really NEC also that wash.PL clean hand
'For employees: You must wash hands.
For the general public: You should really also wash hands.'

The paper is organised as follows: Section 2 reviews some of the core properties of weak necessity modality and introduces the domain restriction approach. It also surveys the morpho-syntactic strategies that languages adopt to encode weak strength. We then situate Afrikaans and Samoan within this typology in Section 3, while also identifying a previously unattested strategy of leaving the distinction between weak and strong necessity unmarked. Section 4 proposes an analysis for this apparent variability in strength. We conclude in Section 5 with a brief discussion of variability across the different dimensions of modal meaning across languages.

2 Background

2.1 Weak Necessity

Weak necessity 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 (4). They entail possibility, while being entailed by their stronger counterpart (Horn, 1972; Rubinstein, 2021; von Fintel & Iatridou, 2008). This entailment relation gives rise to a scalar implicature to the exclusion of the stronger claim. This implicature can be overtly reinforced, cancelled or suspended, as illustrated in (5).

- (4) *You must/should/can stay but you #must/#should/^{ok}can also go.*
- (5) *I ought to help the poor.*
[Reinforcement:] *But I don't have to.*
[Cancellation:] *In fact, I must.*
[Suspension:] *Maybe I have to.*

Building on these properties, Rubinstein (2012, p. 4, no. 4) proposes the working definition in (6), which also serves as the working definition for the research on Afrikaans and Samoan presented

SG = singular, TAM = tense-aspect marker, WNEC = weak necessity, and X = x-marking.

Weak strength in a necessity can then be modelled as resulting from a smaller domain of quantification (von Fintel & Iatridou, 2008, p. 118): “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.” One possible implementation of this idea is in (9), where the domain of quantification depends on a second ordering source. Quantification is thus over the best of the best worlds, as shown in Figure 1.

$$(9) \quad \llbracket (\text{weak necessity modal}) \rrbracket = \lambda a_{\langle s,t \rangle} \cdot \lambda o1_{\langle \langle s,t \rangle, t \rangle} \cdot \lambda o2_{\langle \langle s,t \rangle, t \rangle} \cdot \lambda p_{\langle s,t \rangle} \cdot \\ \forall w' [w' \in \text{BEST}(o2, \text{BEST}(o1, a)) \rightarrow p(w') = 1]$$

The analysis of weak strength through an additional domain restriction is particularly attractive for languages in which weak and strong necessity modal expressions are morpho-syntactically transparently related. We take a closer look at the different strategies that languages adopt for weak necessity in the next section.

2.3 The Typology of Weak Necessity

We review here three morpho-syntactic strategies that languages adopt for weak necessity, even though “a larger and more balanced cross-linguistic investigation of the phenomenon is clearly called for.” (Rubinstein, 2012, p. 5) Under the first strategy, weak necessity is lexically encoded (von Fintel & Iatridou, 2008). Under a second strategy, weak necessity is morphologically marked (von Fintel and Iatridou, 2008, 2023; Vander Klok and Hohaus, 2020). A third strategy is the use of a comparative paraphrase (Rubinstein, 2014). We note with Vander Klok and Hohaus (2020) that languages may adopt multiple of these strategies. Dutch (Indo-European, Germanic) is one such language (see also Hohaus et al., 2023): Weak necessity is lexicalised in the verb *horen*, in (10), but may also be marked morphologically, as in (11). Here, the language borrows morphology that we also find in the consequent of a counterfactual conditional, such as (12), which von Fintel and Iatridou (2023) refer to as x-marking.

- (10) *Je hoort dat zo te doen.* – Dutch –
 you WNEC this so to do
 ‘You should do it this way.’
- (11) a. *Je zou eens Anna Karenina moeten lezen, maar het hoeft niet.*
 you X sometime NAME NEC read but it NEC not
 ‘You should read *Anna Karenina* sometime, but you don’t have to.’
 (von Fintel & Iatridou, 2008, p. 124, no. 31)
- b. *#Je moet Anna Karenina lezen, maar het hoeft niet.*
 you NEC NAME read but it NEC not
 ‘You have to read *Anna Karenina*, but you don’t have to.’
 (von Fintel & Iatridou, 2008, p. 124, no. 32)
- (12) *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.’
 (von Fintel & Iatridou, 2008, p. 124, no. 30)

The use of x-marking appears to be widespread across languages; von Fintel and Iatridou (2008) propose the generalisation in (13).

- (13) The Consequent X-Marking Generalisation:
 “...it is a cross-linguistically stable fact that the meaning of OUGHT can be conveyed with counterfactual morphology on a strong necessity modal.”
 (von Fintel & Iatridou, 2008, p. 126)

The morphological strategy used for weak necessity is however not limited to x-marking (Vander Klok & Hohaus, 2020): Pakiran Javanese (Austronesian, Javanese) has a dedicated particle *-ne*, which weakens the strength of a necessity (but not a possibility). An example is in (14).

- (14) *Wong wong jawa kudu-ne iso ngomong kromo,* – Javanese –
 person person java ROOT.NEC-NE CIRC.POS AV.talk high.speech
terus anak-e rojo yo kudu iso.
 then child-DEF king PRT.YES ROOT.NEC CIRC.POS
 ‘Javanese people ought to be able to speak *Krama*, the Sultan’s son has to.’
 (Vander Klok & Hohaus, 2020, p. 2, no. 2)

Returning to Dutch, an example of the use of a comparative paraphrase is in (15). Note that this is the only strategy for weak necessity in Hebrew (Afro-Asiatic, Semitic; Rubinstein, 2014).

- (15) *Het is beter dat ze niet met hem meegaat.* – Dutch –
 it is better that she not with him with.go
 ‘She better not go with him; she shouldn’t go with him.’

In the next section, we present data that identify a fourth, new strategy: Afrikaans and Samoan may leave the distinction between weak and strong necessity unmarked.

3 Data

In this section, we review the availability of the different strategies for weak necessity in two typologically unrelated languages: Afrikaans, an Indo-European, Germanic language spoken by around 17 million people in South Africa, and Samoan, an Austronesian, Oceanic language spoken by approximately 175,000 speakers on the Samoan archipelago (Simons & Fennig, eds., 2023).⁴

Force and Flavour. In broad strokes, both languages employ a paradigm of modal (auxiliary) verbs that encodes both force and flavour distinctions (de Villiers, 1971, Donaldson, 1993, Erasmus, 2019; Mosel and Hovdhaugen, 1992, Hohaus, 2020), much like English. We focus here on two modal expressions in particular, Afrikaans *moet* and Samoan *tatau*, which we will argue exhibit variable strength. Both encode necessity force, as illustrated in (16) and (17). In relation

⁴Unless otherwise indicated, the data come from elicitation with native speakers relying on translation, acceptability judgment and targeted production tasks (see also Matthewson, 2004, for instance). SW is a heritage speaker and has been working on Afrikaans since 2021. VH has been working on Samoan since 2009. The research presented here underwent ethical review and approval at the University of Manchester.

to flavour, Afrikaans *moet* is of variable acceptability in epistemic contexts and appears to be restricted to root contexts for many of our consultants (but see Erasmus, 2019): Compare the root use in (18) with the epistemic case in (19), for example. Samoan *tatau*, however, is used across flavours, with (20) an example of a root interpretation and (21) targeting an epistemic reading.

- (16) #*Die hond moet buite bly en hy moet ook in=kom.* – Afrikaans –
 the dog NEC outside stay and he NEC also in=come
 ‘The dog must stay outside and he must also come inside.’
- (17) Prompt: “Can both sentences be true at the same time?”
 a. *O Vela e tatau ona i totonu/ fafo o le fale.* – Samoan –
 FOC NAME TAM NEC that PREP inside/ outside of the house
 ‘Vera must be inside/ outside.’
 Comment: “You cannot be in two places at the same time.”
- (18) In England, the law states that when you ride a motorbike:⁵
 a. *Jy moet ’n helm dra.* – Afrikaans –
 you NEC a helmet wear
 ‘You must wear a helmet.’
- (19) In the evening, you see that Bee’s light is on at her house, so you think:⁵
 a. %*Die lig brand, so Bee moet by die huis wees.*
 the light burns so NAME NEC at the house be.INF
 Int. ‘The light is on, so Bee must be at home.’
- (20) Preparing for the theory test for a Samoa driver’s license.
 a. *E tatau ona ‘e ta’u=avanoa i tagata savavali pe ‘a* – Samoan –
 TAM NEC that you make=space to person walk.PL if
liliu i le itu agavale a’o mumū mai le moli lanumumū.
 turn to the side left while red from the light red
 ‘You must give way to pedestrians when turning left at a red light.’
- (21) The policeman is certain that he knows where John and Jodi are hiding.⁶
 a. *E tatau la ona i tua o le ie fa’amalama!*
 TAM NEC then that PREP behind of the cloth window
 ‘They must be behind the curtain then!’

While the above examples express strong necessity, the strength of both modal verbs appears variable. Weak strength thus appears to be go morphologically unmarked in both languages. Before discussing this observation in more detail, we review which of the strategies for weak modal strength (and weak necessity in particular) discussed in Section 2.3 the two languages employ.

⁵From J. Vander Kloek (2022), “Revised Modal Questionnaire for Cross-Linguistic Use” (URL: <https://jozinav.wordpress.com/linguistics/>), accessed 16th April 2023).

⁶From TFS Working Group (2011), “On the Lam,” *Totem Field Storyboards* (URL: <http://totemfieldstoryboards.org/>), accessed 10th August 2019).

Lexicalised Weak Necessity. Afrikaans lexically encodes weak necessity in the quasi-auxiliary verb *behoort* (de Villiers, 1971; Erasmus, 2019), which is restricted to certain more polite or formal registers and, for our consultants, is generally dispreferred. A relevant example is in (22). This auxiliary verb meets the working definition for a weak necessity modal: It patterns with necessity modality with respect to the conjunction diagnostic, as shown below in (23), and entails possibility modality, as elicited by means of (24).

- (22) *Die eersteling behoort eerste te trou.* – Afrikaans –
 the firstborn WNEC first to marry
 ‘The firstborn should marry first.’⁵
- (23) #*Die hond behoort buite te bly en hy behoort ook in=te=kom.*
 the dog WNEC outside to stay and he WNEC also in=to=come
 ‘The dog should stay outside and he should come inside.’
- (24) Assume that you have asked how to travel to the shops, and I tell you the below.
 How would you then answer the question that follows?
- a. *As jy winkels toe gaan behoort jy die kar te neem.*
 if you shops to go WNEC you the car to take
 ‘If you go to the shops, you should take the car.’
- b. *Kan jy die kar neem?* c. *Ja.*
 POS you the car take yes
 ‘Can you take the car?’ ‘Yes.’

Evidence in favour of a weak interpretation of *behoort* comes from the implicature it may give rise to, to the exclusion of a stronger reading with *moet* or the negative polarity *hoef*. This implicature can be reinforced, cancelled and suspended, as shown in (25).

- (25) a. *Behoort ek met haar huiswerk te help? Ek weet ek hoef nie.*
 WNEC I with her homework to help I know I NEC NEG
 ‘Should I help with her homework? I know I don’t have to.’
- b. *Jy behoort eintlik ‘n elektriese toets te doen,*
 you WNEC actually a electric test to do
om die waarheid te sê, jy moet.
 in.order the truth to say you NEC
 ‘You should actually do an electrical test, to tell you the truth, you must.’
- c. *Behoort ek hierdie op my belastingopgawe te verklaar? Dalk moet ek?*
 WNEC I this on my tax.return to declare perhaps NEC I
 ‘Should I declare this on my tax return? Perhaps I have to?’

Turning to the paradigm of modal verbs in Samoan, we have been unable to identify a lexicalised distinction between weak and strong necessity (see also Mosel and Hovdhaugen, 1992; Hohaus, 2020). Prompts with a weak necessity modal expression, such as (26), consistently receive a translation with *tatau*, as do prompts with strong necessity modality.

- (26) Prompt: “You should check your tyres before you start driving, but it’s not a law.”
- a. *E tatau ona siaki muamua pa’u o lou ta’avale* – Samoan –
 TAM NEC that check first tire of your car
a’o le’i alu ‘ese. E le=’o se tulafono.
 while not.yet go away TAM not=FOC a law
 ‘You should check the tires of your car before you leave. It’s not a law.’
 Comment: “We don’t have a separate word for *should*; it is always *tatau*.”

We will revisit this observation below, once we have reviewed the remaining typologically attested patterns for weak necessity.

Morphological Marking. Neither Afrikaans nor Samoan fall under the Consequent X-Marking Generalisation in (13) above. In Afrikaans, the consequent of a counterfactual conditional such as (27) features the modal auxiliary verb *sou* (see also Erasmus, 2019, pp. 608-614), the past tense form of *sal* ‘shall’. As a strategy to derive weak necessity, consequent x-marking however appears to be unavailable for our consultants. Relevant configurations result in a strong counterfactual, rather than weak, necessity, as shown in (28) and (29), an interpretation which von Fintel and Iatridou (2023, p. 28) refer to as the exo-x reading.

- (27) The dads are discussing their weeks and how much work they ended up doing.
 Jack laughs and says: “Well, . . .”
- a. *As ek ’n Prokereer was, dan sou ek meer geld verdien.* – Afrikaans –
 if I a lawyer was then X I more money earn
 ‘If I were a lawyer, I would earn more money.’
- (28) *Ek sou Londen toe moes gaan vandag om my paspoort te kry. . .*
 I X NAME to NEC.PST go today in.order my passport to get
 ‘I would have had to go to London today in order to get my passport. . .’
 Int. ‘I should have gone to London today. . .’
- (29) *Jan sou hierdie boeke moet begin lees.*
 NAME X these books NEC begin read
 ‘Jan would have to begin reading these books.’
 (Conradie, 2003, p. 82, no. 39-a)
 Comment: “It is a bit unusual.”

In Samoan, counterfactuality goes unmarked in the consequent of a conditional. Counterfactuality is instead indicated in the antecedent in the choice of complementiser (see also Mosel and Hovdhaugen, 1992; Wierzbicka, 1997), between realis ‘*afai* ‘if’ and irrealis ‘*ana*, as shown in (30).

- (30) a. ‘*Ae afai ‘o se pese silou, ia, ‘ua uosi.* – Samoan –
 but if FOC a song slow well TAM.INCH waltz
 ‘But if it was slow song, they started dancing a waltz.’
 (Mosel & Hovdhaugen, 1992, p. 654, no. 16.3, our glosses)

- b. *Ana ou pasi ua fiagofie foi ona maua sau galuega.*
 if.IRR I pass TAM.INCH easy also that get my work
 ‘If I had passed the exam, it would also have been easy to get work.’
 (Mosel & Hovdhaugen, 1992, p. 657, no. 16.21, our glosses)

The Consequent X-Marking Generalisation thus does not hold for independent reasons in Samoan, unless we allow for x-marking to be null.

Comparative Paraphrases. In both languages, consultants frequently resort to comparative paraphrases in the elicitation of weak necessity (see also Hohaus, 2021). Relevant examples are in (31), from Afrikaans, and (32) to (33), for Samoan.

- (31) a. Prompt: “You should do it this way, but it’s up to you.”
 b. *Dis beter om dit so te doen, maar jy kan dit ook so doen.* – Afrikaans –
 it.is better in.order it so to do but you can it also so do
 ‘It’s better to do it like so, but you can also do it like so.’
- (32) Your friend Sina is visiting, but one of your sisters is about to return;
 she doesn’t like her. You advise Sina: – Samoan –
 a. *E sili ona e alu loa, ae a leai, e leai se mea o iai.*
 TAM great that you go immediately but when NEG TAM NEG a thing REL have
 ‘It would be best if you leave but, if not, it’s not a thing.’
- (33) Too much to do, not enough time! Abella is hungry, and you need to wash your clothes.
 a. *E sili atu ona fafaga e Me le pepe nai lō le tā o lāvalava.*
 TAM great more that feed ERG NAME the baby from COMP the wash of cloths
 ‘It would be better for Me to feed the baby than do the laundry.’

Briefly summarising the empirical findings so far, of the different strategies for weak modal strength attested, Afrikaans has the lexicalised weak necessity modal verb *behoort*, but does not obey the Consequent X-Marking Generalisation. Samoan may not fall under this generalisation for independent reasons. Comparative paraphrases are attested in both languages. Both languages however also seem to allow for weak necessity to be unmarked, as we show next.

Variable Strength. Afrikaans *moet* is not only compatible with strong interpretations, but also with weak, of which the translation of (34) is suggestive. Recall however also (2) from Section 1, which features both the weak and strong interpretation. Like its English cognate, Afrikaans *moet* is acceptable in contexts where the prejacent describes the only possible course of action, as in (35-a,b,c). Unlike its English cognate (see also von Fintel and Iatridou, 2008), however, it is also acceptable in contexts where multiple other alternatives are available, as shown in (36-d,e,f).

- (34) *Jy moet nou daardie oproep maak.* – Afrikaans –
 you NEC now that call make
 ‘You should now make that call.’⁷

⁷J. Conradie (2020), “Root Semantics”, *Taalportal.org*, no. 1-a (URL: <https://www.taalportaal.org/taalportaal/topic/pid/topic-14857881438688606>, accessed 27th October 2023).

- (35) There are three ways to get to Manchester— the back routes, the M6, and through Reading.
- a. All other roads are closed today, however.
 - b. *If you go to Manchester, you **must** use the M6.*
 - c. *Jy **moet** die M6 gebruik, omdat die ander paaie toe is.*
you NEC the NAME use because the other roads closed are
'You have to use the M6 because the other roads are closed.'
 - d. Bee says that the route using the M6 is the best.
 - e. *#If you go to Manchester, you **must** use the M6.*
 - f. *As jy na Manchester toe gaan, **moet** jy die M6 gebruik.*
if you to NAME to go NEC you the NAME use
'If you go to Manchester, you should use the M6.'

The strength variability of Afrikaans *moet* also shows up in (36) to (37), where the weak reading appears to exclude a strong necessity. This meaning can then be reinforced, cancelled or suspended, using a strong reading of *moet* in the continuations, where grammatically possible. The strength in this case appears to be governed exclusively by context.

- (36) It's your last day at work before your leave, and there's a lot left to do. You catch sight of your calendar and realise the summer party is this afternoon. Talking to your colleague:
- a. *Ek **moet** nog na die partyjie toe gaan!*
I NEC still to the party to go
'I still should/have to go to the party!'
- (37)
- a. It's written in pencil on your calendar, so you know you did not confirm.
*Maar eintlik **hoef** ek nie te gaan nie.*
but actually NEC I NEG_{1/2} to go NEG_{2/2}
'But actually, I don't have to.'
 - b. It's written on your calendar by your boss, in pen, so you realise you have to go.
*Streng gesproke, **moet** ek gaan.*
strict PST.speak NEC I go
'Strictly speaking, I have to go.'
 - c. You aren't sure whether staff have to be there, or whether it was left up to you.
You think perhaps it is compulsory.
*Miskien **moet** ek maar gaan.*
maybe NEC I but go
'Perhaps I have to go.'

Recall from (3) and (26) above that the Samoan necessity modal verb *tatau* also appears to be compatible with both strong and weak interpretations. Another example is in (38), elicited using the contexts from (32) above. Just like in Afrikaans, under the weak reading, the modal verb is compatible with context that set up multiple viable alternatives, such as in (39). Here, a win for Samoa is not certain, and another outcome is a possibility. The strong interpretation arises, for instance, in contexts where the truth of the prejacent is guaranteed, such as (40).

- (38) *E tatau ona e alu loa,* – Samoan –
 TAM NEC that you go immediately
ae a leai, e leai se mea o iai.
 but when NEG TAM NEG a thing REL have
 ‘You should leave, but, if not, it’s not a thing.’
- (39) Last night, Tala briefly switched on the television to check on the rugby game. The All Blacks have a narrow lead on Manu Samoa when Tala switches the television off again. However, he is confident that Manu Samoa would still win. His prediction:
 a. *E tatau ona manumalo le Manu Samoa i le taaloga.*
 TAM NEC that win the NAME at the game
 ‘Manu Samoa should win the game.’
- (40) You play a game where you have to guess under which cup the marble is. After the cups have been shuffled, you check the blue and the yellow cup, and it’s not there.
 a. *E tatau ona iai le mapu i lalo o le ipuinu lanumūmū.*
 TAM NEC that be.there the marble at below of the cup.drink colour.green
 ‘The marble must be under the green cup.’

This context-dependent variability in strength also surfaces in the continuations to (41) in (42), which was elicited using the contexts already reported in (36) and (37) above for Afrikaans.

- (41) *E tatau ona ou alu i le pātī.*
 TAM NEC that I go to the party
 ‘I should go to the party.’
- (42) a. *Ae e lē tatau iā te a’u.*
 but TAM not NEC to PRN.1.SG
 ‘But I don’t have to.’
 b. *O le upu moni e tatau iā te a’u.*
 FOC the word true TAM NEC to me
 ‘Strictly speaking, I have to.’
 c. *Atonu e tatau iā te a’u.*
 maybe TAM NEC to me
 ‘Maybe it is necessary that I go.’

We take the above data to be incompatible with a characterisation of Afrikaans *moet* and Samoan *tatau* as lexically encoding strong necessity. Rather, the observed variability in strength suggests that weak modal strength may go unmarked in both languages and depend on the context. In the typology of weak necessity summarised in Section 2.3 above, this is a previously unattested strategy. Our focus in the next section is on an analysis of this variability.

4 Analysis

Building on Weingartz (2022), we propose here an analysis of Afrikaans *moet* and Samoan *tatau* under which both modal expressions lexically encode weak necessity. More formally, within the domain restriction approach to weak necessity outlined in Section 2.1, both lexically allow for a second ordering source, as sketched in (43). To account for the observed variability in strength, we propose that this second ordering source may however remain empty, if context does not make available any additional considerations that qualify the necessity. (Note that the idea of empty ordering sources is not novel; see Kratzer (1981, 2012), Sæbø (1985), Peterson (2010), and Miho et al. (2019) for relevant discussion relating to the first ordering source.)

$$(43) \quad \llbracket \textit{moet}_{\text{Afrikaans}} \rrbracket = \llbracket \textit{tatau}_{\text{Samoan}} \rrbracket = \lambda a_{\langle s,t \rangle} \cdot \lambda o1_{\langle \langle s,t \rangle, t \rangle} \cdot \lambda o2_{\langle \langle s,t \rangle, t \rangle} \cdot \lambda p_{\langle s,t \rangle} \cdot \forall w' [w' \in \text{BEST}(o2, \text{BEST}(o1, a)) \rightarrow p(w') = 1]$$

If the secondary ordering source is an empty set in the world of evaluation, the weak necessity claim ends up equivalent to a strong necessity claim with a single ordering source (see also Rubinstein, 2012, pp. 39–48 and Vander Klok and Hohaus, 2020, pp. 10–12 for discussion): If the secondary ordering source is an empty set, the best accessible worlds in relation to the first ordering source will not undergo another round of ranking and selection, that is, for any accessibility relation a , ordering source $o1$ and empty ordering source $o2$, $\text{BEST}(a, o1) = \text{BEST}(o2, \text{BEST}(a, o1))$. The set of favoured worlds, the domain of quantification, remains unchanged.

Under this view, English weak necessity *should* and *ought* differ from their Afrikaans and Samoan counterparts in that they do not allow for the secondary ordering to be empty. We capture this as a presupposition in (44) below, where we set aside any additional restrictions relating to the accessibility relation and ordering sources.

$$(44) \quad \llbracket \textit{should}_{\text{English}} \rrbracket = \lambda a : a \in D_{\langle s,t \rangle} \cdot \lambda o1 : o1 \in D_{\langle \langle s,t \rangle, t \rangle} \cdot \lambda o2 : o2 \in D_{\langle \langle s,t \rangle, t \rangle} \ \& \ \{q \in D_{\langle s,t \rangle} : o2(q) = 1\} \neq \emptyset \cdot \lambda p : p \in D_{\langle s,t \rangle} \cdot \forall w' [w' \in \text{BEST}(o2, \text{BEST}(o1, a)) \rightarrow p(w') = 1]$$

Revisiting the typological classification of Afrikaans and Samoan in relation to the strategies adopted for weak necessity, let us point out that both languages lexicalise weak strength under the analysis proposed. While unmarked strength may be a distinct morpho-syntactic strategy, it does not translate to a distinct semantic strategy in Afrikaans and Samoan. The observed variability, we propose, is a result of the flexible interaction with context.

5 Concluding Remarks

The paper identified a new pattern in the morpho-syntactic typology of weak necessity: Languages may leave the distinction between weak and strong necessity morphologically unmarked. Afrikaans *moet* and Samoan *tatau* allow for such null marking and exhibit variable strength. Under the analysis proposed here, the variability can be captured under a uniform analysis, under which the two expressions encode weak necessity modality and are interpreted relative to a secondary ordering source. This set may however also be the empty set, resulting in a strong interpretation.

While null marking is a previously unattested strategy, no additional new technology is required for its formal analysis under the domain-restriction approach to weak necessity (Rubinstein, 2012; von Stechow & Iatridou, 2008). Both languages may also additionally resort to a comparative paraphrase, but interestingly do not necessarily fall under the Consequent X-Marking Generalisation (von Stechow & Iatridou, 2008, 2023). In Afrikaans, consequent x-marking results in counterfactual rather than weak necessity. In Samoan, counterfactuality is marked on the complementiser of the conditional antecedent only and may not be in the scope of the generalisation. It may be worth exploring however in how far x-marking may be characterised as the null marking we have described here for weak necessity.

Returning to the triad of force, flavour and strength that characterises modal meaning, our data can be taken to complement a productive line of cross-linguistic research in the past fifteen years that has explored the variability of force and flavour in modal expressions across languages. Flavour may be variable in the presence of constant force, as in English (45), but we also find the opposite pattern (prominently, Rullmann et al., 2008; Deal, 2011; Bochnak, 2015). St’át’imcets (Salishan, Central Salish; Canada), for instance, has variable force, but constant flavour, as illustrated in (46).

(45) *John may_{root} not watch TV, but he may_{epistemic} be watching it anyway.*
(Hacquard, 2010, p. 81)

- (46) a. *wa7 k’a séna7 qwenúxw* – St’át’imcets –
IMPF INFER COUNTER sick
‘He may be sick. (Maybe that’s why he’s not here.)’
(Rullmann et al., 2008, p. 321, no. 5-d)
- b. You have a headache that won’t go away, so you go to the doctor.
All the tests show negative. There is nothing wrong, so it must just be tension.
- c. *nilh k’a lh(el)-(t)-en-s-wá (7)-(a) ptinus-em-sút*
FOC INFER from-DET-1SG.POSS-NOM-IMPF-DET think-MID-OOC
‘It must be from my worrying.’
(Rullmann et al., 2008, p. 321, no. 5-c)

Based on the view from Afrikaans and Samoan, this variability can be found across all three dimensions of modal meaning and extends to strength. In all cases, however, the observed variability derives systematically from an underlyingly uniform semantics.

References

- Bochnak, M. R. (2015). Variable force modality in Washo. *Proceedings of the Annual Meeting of the North Eastern Linguistic Society (NELS) 45*, 105–114.
- Conradie, C. J. (2003). Isomorphism and language change. In B. J. Blake, K. Burridge, & J. Taylor (Eds.), *Historical Linguistics 2001: Selected Papers from the 15th International Conference on Historical Linguistics (ICHL)* (pp. 71–86). John Benjamins.
- Deal, A. (2011). Modals without scales. *Language*, 87(3), 559–585.

- de Villiers, M. (1971). *Die grammatika van tyd en modaliteit*. Balkema.
- Donaldson, B. C. (1993). *A grammar of Afrikaans*. De Gruyter.
- Erasmus, C. J. (2019). *Op soek na 'n omvattende tipologiese beskrywing van modaliteit: Die Afrikaanse modale sisteem* [Doctoral dissertation, Noordwes-Universiteit].
- Hacquard, V. (2010). On the event relativity of modal auxiliaries. *Natural Language Semantics*, 18(1), 79–114.
- Hohaus, V. (2020). *Flavours of Samoan modality: The split between epistemic and root* [Manuscript, The University of Manchester].
- Hohaus, V. (2021). Gradability and modality: A case study from Samoan. In L. Clemens & D. Massam (Eds.), *Polynesian syntax and its interfaces* (pp. 11–35). Oxford University Press.
- Hohaus, V., Weingartz, S., & Boomars, M. (2023). Variable strength in necessity modality: A case of variation between Afrikaans, Dutch, English, and German. In L. Matthewson, E. Csipak, M. R. Bochnak, M. Morzycki, & D. Reisinger (Eds.), *Festschrift for Hotze Rullmann* (to appear). University of British Columbia Working Papers in Linguistics (UBCWPL).
- Horn, L. R. (1972). *On the semantic properties of logical operators in English* [Doctoral dissertation, University of California].
- Klecha, P. (2014). *Bridging the divide: Scalarity and modality* [Doctoral dissertation, University of Chicago].
- Kratzer, A. (1977). What *must* and *can* must and can mean. *Linguistics and Philosophy*, 1(5), 337–355.
- Kratzer, A. (1981). The notional category of modality. In H.-J. Eikmeyer & H. Rieser (Eds.), *Words, worlds, and contexts: New approaches in word semantics* (pp. 38–74). De Gruyter.
- Kratzer, A. (1991). Modality. In A. von Stechow & D. Wunderlich (Eds.), *Semantik: Ein internationales Handbuch der zeitgenössischen Forschung* (pp. 639–650). De Gruyter.
- Kratzer, A. (2012). *Modals and conditionals*. Oxford University Press.
- Lassiter, D. (2017). *Graded modality: Qualitative and quantitative perspectives*. Oxford University Press.
- Matthewson, L. (2004). On the methodology of semantic fieldwork. *International Journal of American Linguistics*, 70(4), 369–451.
- Matthewson, L. (2016). Modality. In M. Aloni & P. Dekker (Eds.), *Cambridge handbook of formal semantics* (pp. 726–775). Cambridge University Press.
- Miho, T., Bhadra, D., & Fălăuş, A. (2019). Epistemic modals, deduction, and factivity: New insights from the epistemic future. *Proceedings of Semantics and Linguistic Theory (SALT)* 29, 351–370.
- Mosel, U., & Hovdhaugen, E. (1992). *Samoan reference grammar*. Scandinavian University Press.
- Peterson, T. R. (2010). The role of the ordering source in Gitksan epistemic modals. *Proceedings of Semantics of Under-Represented Languages of the Americas (SULA)* 6, 171–192.
- Portner, P., & Rubinstein, A. (2016). Extreme and non-extreme deontic modals. In N. Charlow & M. Chrisman (Eds.), *Deontic modality* (pp. 256–282). Oxford University Press.
- Rubinstein, A. (2012). *Roots of modality* [Doctoral dissertation, University of Massachusetts, Amherst].
- Rubinstein, A. (2014). On necessity and comparison. *Pacific Philosophical Quarterly*, 95(4), 512–554.

- Rubinstein, A. (2021). Weak necessity. In D. Gutzmann, L. Matthewson, C. Meier, H. Rullmann, & T. E. Zimmermann (Eds.), *The Wiley Blackwell Companion to Semantics* (pp. 1–44, Vol. 4). Wiley-Blackwell.
- Rullmann, H., Matthewson, L., & Davis, H. (2008). Modals as distributive indefinites. *Natural Language Semantics*, 16(4), 317–357.
- Sæbø, K. J. (1985). *Notwendige Bedingungen im Deutschen: Zur Semantik modalisierter Sätze* [Doctoral dissertation, Universitetet i Oslo].
- Simons, G. F., & Fennig, eds., C. D. (2023). *Ethnologue: Languages of the world* (26th ed.). SIL International. <http://www.ethnologue.com>
- Sloman, A. (1970). *Ought and better*. *Mind*, 79(315), 385–394.
- Vander Klok, J., & Hohaus, V. (2020). Weak necessity without weak possibility: The composition of modal strength distinctions in Javanese. *Semantics and Pragmatics*, 13(12), 1–50.
- von Stechow, P., & Iatridou, S. (2008). How to say *ought* in Foreign: The composition of weak necessity modals. In J. Guéron & J. Lecarme (Eds.), *Time and modality* (pp. 115–141). Springer.
- von Stechow, P., & Iatridou, S. (2023). Prolegomena to a theory of x-marking. *Linguistics and Philosophy*, online first, 1–44.
- Weingartz, S. (2022). *Modal necessity in Ndebele (Bantu; Zimbabwe) and Afrikaans (Germanic; South Africa)* [MA dissertation]. The University of Manchester.
- Wierzbicka, A. (1997). Conditionals and counterfactuals: Conceptual primitives and linguistics universals. In A. Athanasiadou & R. Dirven (Eds.), *On conditionals again* (pp. 15–60). John Benjamins.

Reduplicated Distributivity in Mandinka¹

Ousmane Cisse — Boston University

Elizabeth Coppock — Boston University

Abstract. Reduplication is commonly exhibited by markers of distributivity. Although distributivity markers can either mark the key (as determiner *each* does, as in *each child saw a lion*) or the share (as with adnominal *each*, as in *the boys saw a lion each*), it has been conjectured that distributivity markers formed through reduplication are always markers of the share, rather than the key. Here we discuss a case that challenges but ultimately vindicates this conjecture. In Mandinka (spoken in Senegambia), reduplicating a nominal with interposition of the morpheme *-woo-* gives rise to a distributive reading. We investigated the semantics of the X-woo-X construction and found that it behaves as a key-marker, but also as a share-marker. We take these findings to support an analysis on which X-woo-X signals ‘simultaneous distributivity’, simultaneously marking both key and share.

Keywords: Reduplication, (simultaneous) distributivity, Mandinka, exhaustivity, share marker, key marker.

1 Introduction

1.1 Gil’s conjecture

This paper discusses a reduplication-based strategy for marking distributivity in Mandinka, a Mande language spoken primarily in Senegal (and the first author’s native language). Here are several examples of this construction, which we call ‘X-woo-X’:

- (1) **Musu-woo-musu** ye kini taboo noo le [Mandinka]
woman-DIST-woman PRED rice cooking know PERF
‘Each woman knows how to cook rice.’
- (2) Fode ye siise-e **kili-woo-kili** samba le.
Fode PRED chicken egg-DIST-egg carry PERF
‘Fode carried each chicken egg’
- (3) Binta ye mangu saamu **kiliq-oo-kiliq** saq ne
Binta PRED mango pile one-DIST-one buy PERF
‘Binta bought the mangoes one by one / each mango.’

Along with interpolation of the element *-woo-* (which is also used as the demonstrative ‘this’), this construction involves reduplication, either of a noun, as in (1) or a numeral, as in (3). As shown by the gloss, X-woo-X can be translated as ‘each X’, and generally contributes universal force.

¹We would like to thank the audience at TripleA 10 in Potsdam for excellent discussion, especially Jérémy Pasquereau, Jakob Maché and Malte Zimmermann.

It is not uncommon cross-linguistically that reduplication of nouns gives rise to an interpretation that is paraphrased with ‘each’ or ‘every’. Moravcsik (1978) lists examples including the following (see paper for references):

- (4) a. *òsòòsè* ‘every week’ (cf. *òsè* ‘week’) [Yorùbá]
 alalé ‘every enemy’ (cf. *alé* ‘enemy’)
 b. *araw^cáraw* ‘every day’ (cf. *araw* ‘day’) [Tagalog]
 c. *renren* ‘everybody’ (cf. *ren* ‘man’) [Mandarin]

About these types of cases, Gil (1995, 335) writes, “Although at first blush reduplication appears to bear the denotation of distributive-key universal quantifier, closer inspection reveals subtle distinctions.”

In his intriguing comment, Gil invokes the notion of ‘distributive key’; let us unpack that before addressing Gil’s view on reduplicated nouns. The notion of ‘(distributive) key’ can be explained in contradistinction to the notion of ‘share’ using adnominal *each* in English, as in:

- (5) The kids carried **five balloons** each.
 key = the kids; share = five balloons

This sentence expresses a distributive relation where for each of the kids, there are five balloons. There is universal quantification over the kids, taking scope over existential quantification related to ‘5 balloons’. Generally, a distributive relation involves universal quantification taking scope over existential quantification, as schematized on the lefthand side in Figure 1. The ‘key’ is the set restricting the universal quantifier (or the noun phrase corresponding to it), and the ‘share’ is the set restricting the existential quantifier (or the corresponding noun phrase). Thus, in this example, *the kids* is the key and *five balloons* is the share.² To say that reduplicated nouns appear at first blush to be distributive-key universal quantifiers is to say that they appear at first blush to associate with the key in a distributive relation, bearing universal force.

The “subtle distinctions” that Gil alludes to have to do with event-key readings of distributivity markers. These can be illustrated with examples from Korean and Telugu. Korean *-ssik* behaves much like binominal *each*, attaching to the share in a distributive relation whose key is determined by a noun phrase found elsewhere in the sentence (Choe, 1987):

- (6) ai-tul-i **phwungsen-hana-ssik-ul** sa-ess-ta [Korean]
 child-PL-NOM balloon-one-SSIK-ACC bought
 ‘The children bought a balloon each.’

Unlike English *each*, however, Korean *-ssik* has so-called ‘event key’ readings where there is no nominal in the sentence that serves as the key, and the set universally quantified over appears to be a set of events described by the verb (Choe, 1987, 52):

- (7) na-nun **phwung-hana-ssik-ul** sa-ess-ta [Korean]
 I-TOP balloon-one-SSIK-ACC bought
 ‘I bought one balloon each time’

²A helpful mnemonic for remembering which is the share and which is the key is the template ‘SHARE per KEY’ (Gil, 2013) – in this case, there are five balloons per kid.

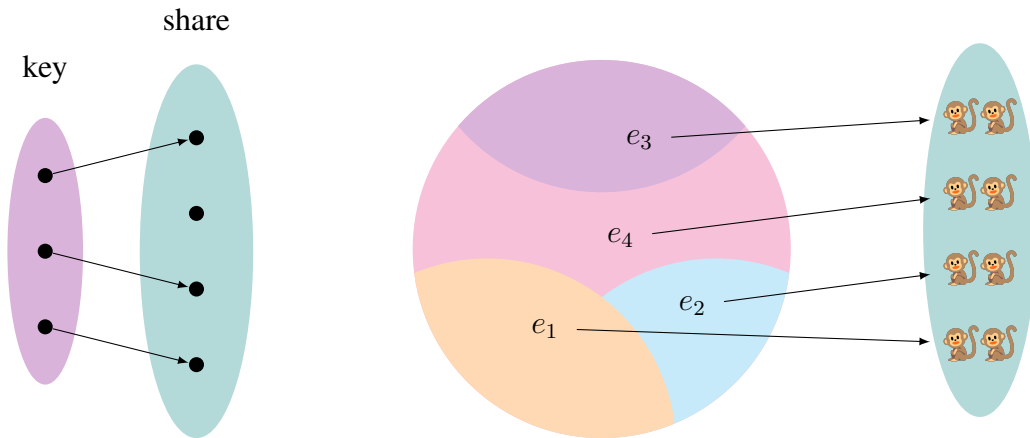


Figure 1: A distributive relation (left). Distributing ‘two monkeys’ over subevents (right).

Reduplicated numerals in Telugu also have event key readings, as Balusu (2006) points out. (8) is ambiguous between a participant key reading and two different types of event key readings.

- (8) ii pilla-lu **renDu renDu** kootu-lu-ni cuus-ee-ru [Telugu]
 these kid-PL 2 2 monkey-PL-ACC see-PAST-3PL
 lit. ‘These kids saw 2 2 monkeys’
 ... each saw 2 monkeys. Participant key
 ... saw 2 monkeys each time. Temporal key
 ... saw 2 monkeys in each location. Spatial key

The participant key reading can be paraphrased ‘every kid saw two monkeys’. One of the event key readings is temporal and the other is spatial. On the ‘temporal key’ reading, the kids saw two monkeys at each time. On the spatial key reading, the kids saw two monkeys in each location.

With some uses of reduplicated numerals in Telugu, event key readings are the only sorts of readings available. In neither of the following examples is there a plural definite NP that would work as an indicator of what the participant key would be:

- (9) Raamu **rendu renDu** kootu-lu-ni cuus-ee-Du [Telugu]
 Ram 2 2 monkey-PL-ACC see-PAST-2PL
 lit. ‘Ram saw 2 2 monkeys ...’
 ... each time. Temporal key
 ... in each location. Spatial key
- (10) **renDu renDu** kootu-lu egir-i-nyiyi
 2 2 monkey-PL jump-PAST-3PL
 ‘2 monkeys jumped in each time/location’

In both of these cases, the reduplicated numeral is associated with the share in an event key distributive relation. Balusu envisions an analysis of event-key readings with an event that is divided up into sub-events, each of which is associated with a pair of monkeys. An example of such a

state of affairs is depicted in Figure 1 (right), where the circle represents an event and the various subregions of it represent sub-events of it.

Back to reduplicated *nouns*: Considering Hebrew examples like the following, Gil (1995) draws a parallel between reduplicated nouns and reduplicated numerals.

- (11) a. haʔanašim saḥvu mizvada yom yom [Hebrew]
 the.man.PL.M carry.PAST.3PL suitcase day day
- b. haʔanašim saḥvu **mizvada mizvada**
 the.man.PL.M carry.PAST.3PL suitcase suitcase
- c. haʔanašim saḥvu et ha-mizvadot **aḥat aḥat**
 the.man.PL.M carry.PAST.3PL ACC the-suitcase.PL:F one.F one.F
- d. haʔanašim saḥvu et ha-mizvadot **šaloš šaloš**
 the.man.PL.M carry.PAST.3PL ACC the-suitcase.PL:F three.F three.F

Examples (11a) and (11b) involve reduplicated nouns; (11c) and (11d) involve reduplicated numerals. Gil points out that (11b), with ‘suitcase suitcase’ is nearly synonymous with (11c), ‘carried the suitcases one one’. He takes it to be uncontroversial that in (11c), the reduplicated numeral marks the share in an event key distributive relation. Based on the synonymy of (11b) and (11c), Gil suggests that (11c) is really a case of share marking; in other words, the reduplication is marking the share in an event key distributive relation. He wonders whether this pattern might be universal (p. 336):

From an iconic perspective, it is of course more natural for reduplication to mark distributive shares than distributive keys; however, it is also natural for reduplication to express the notion of universal quantification. Whether there exist *bona fide* instances of reduplication with the interpretation of distributive key universal quantifier must remain open for the future investigation.

The idea that it is more natural for reduplication to mark distributive shares than distributive keys raises the question of whether their doing so is a linguistic universal. Let us define ‘Gil’s conjecture’ as follows:

- (12) **Gil’s conjecture:** Distributivity markers that are reduplicated (numerals or nouns) always mark the share in a distributive relation.

Gil does not state this conjecture directly, but we are nevertheless naming it after him.

The work we are reporting on today provides some support for Gil’s conjecture, albeit in a slightly nuanced way. Sometimes distributivity markers do double-duty, simultaneously marking keys and shares. This phenomenon is known as **simultaneous distributivity** (see Henderson 2019 on Comox-Sliammon and Kuhn & Aristodemo 2017 on French Sign Language). We argue that the Mandinka X-woo-X construction exhibits simultaneous distributivity in this sense, and is thus a hybrid between a share-marker and a key-marker. If so, then there *is* a “bona fide instance of a reduplicated distributivity marker that is interpreted as a distributive key universal quantifier”, and yet Gil’s conjecture may still be universal.



Figure 2: The distribution of Mandinka in Senegambia and its surrounding areas. Map created by Ousmane Cisse in 2022 using Global Mapper; data from The Joshua Project (2022), who credits the Bethany World Prayer Center.

1.2 Mandinka study: General methodology

We'll be focusing on Mandinka as spoken in Senegal, The Gambia, and Guinea Bissau. The number of speakers was estimated at less than 1 million in Senegal in 2017, but it is growing. Mandinka can broadly be classified as a Niger-Congo language, in the Mande subfamily. Alternative names include Mandingue, the local French name, and Socé, the local Wolof name.

For the current study, we collected data in two phases, each characterized by different participant groups and distinct interview methods. Phase I involves ten native speakers of Mandinka from Ziguinchor, comprising five men and five women, with an age range spanning from 20 to over 50 years. The interviews in Phase I were conducted via WhatsApp video conference calls, with participants grouped in pairs or trios. Group interviews open up the possibility that speakers will disagree, discuss their disagreements, and arrive at a consensus, thereby potentially giving an indication of how the observed variation could be weighted in favor of one option or another. Phase II used individual interviews through Zoom video calls, rather than group interviews, because the experimental design involved many variable combinations. We interviewed 12 different

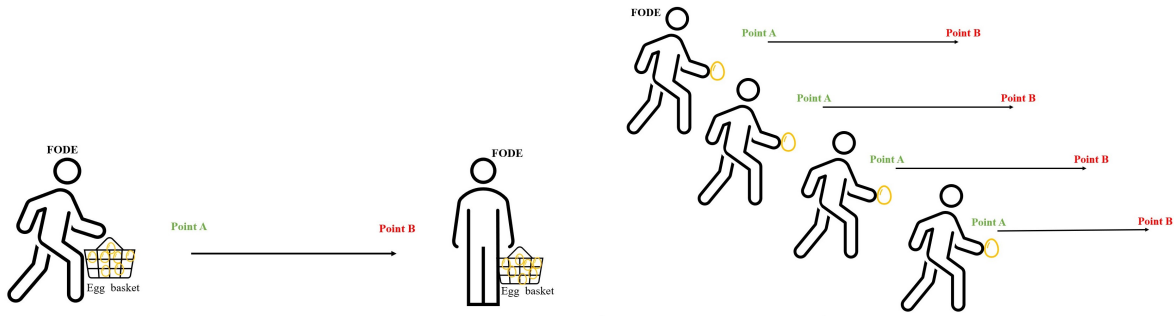


Figure 3: All-at-once scenario (left); one-by-one scenario (right)

native speakers of Mandinka (nine men and three women), also from Ziguinchor, in the same age range as in Phase I.

2 One-by-one effect

We will be establishing two generalizations about the semantics of X-woo-X constructions, starting with the **one-by-one effect**. Suppose that in the X-woo-X construction, X is the distributive share. Then, the sentence involves an event key. Hence there are multiple subevents of the one being described, one per instance of X. Based on this, we predict that X-woo-X should be more felicitous as a way of describing scenarios where the X's are affected one by one, rather than all at once.

With Phase I participants, we collected acceptability judgments relative to the two displays shown in Figure 3. On the lefthand panel of Figure 3, Fode is carrying his eggs all at once from Point A to Point B. We label this the **all-at-once scenario**. On the righthand panel, he takes them one by one. We label this the **one-by-one scenario**.

Relative to these two scenarios, we asked for acceptability judgments on three sentences, one with X-woo-X, one with a definite plural [DEF PL], and one with 'all' [ALL]:

- (13) Fode ye siise-e **kili-woo-kili** samba le.
 Fode PRED chicken egg-DIST-egg carry PERF
 'Fode carried each chicken egg' [X-woo-X]
- (14) Fode ye siise-e **kilo-o-lu** samba le.
 Fode PRED chicken egg-DEF-PL carry PERF
 'Fode carried the chicken eggs' [DEF PL]
- (15) Fode ye siise-e **kilo-o-lu bee** samba le.
 Fode PRED chicken egg-DEF-PL all carry PERF
 'Fode carried all the chicken eggs' [ALL]

Hence all of the sentences involved some way or another of expressing a universal generalization. Participants were asked for acceptability judgements on all three sentences relative to each scenario. We also asked participants which sentence was best given the all-at-once scenario, and which sentence was best given the one-by-one scenario.

	All-at-once scenario	One-by-one scenario
Ex. (13) [X-woo-X]	Infelicitous (unless different kinds)	Good; best choice for scenario
Ex. (14) [DEF PL]	Good	Infelicitous
Ex. (15) [ALL]	Good; best choice for scenario	Infelicitous

Table 1: Acceptability judgments on three sentences relative to one-by-one vs. all-at-once scenarios (Phase I participants)

The participants agreed that the sentences with the definite plural and universal quantifier were acceptable in the all-at-once context but not in the one-by-one scenario. Conversely, the X-woo-X construction was mostly considered unacceptable with the all-at-once scenario. However, one participant raised the point that the sentence could be acceptable if different kinds of eggs are involved. This insight was collectively acknowledged and accepted by all participants.

Furthermore, the participants unanimously concurred that, with the all-at-once scenario, the sentence with the universal quantifier was the most preferred one, although the other sentences were also acceptable. For the one-by-one scenario, we found that the X-woo-X construction was the best way of describing it.

Furthermore, with Phase II participants, we asked for an explanation of the difference in meaning between X-woo-X and sentences involving *bee* ‘all’ with a definite plural, vis-a-vis two scenarios. The sentences were as follows:

(16) *ŋa m baamaa la kitaabu-woo-kitaabu jindi duuma*
 1.SG my mother GEN book-DISTR-book carry down
 ‘I carried down each one of my mother’s books.’

(17) *ŋa m baamaa la kitaabo-o-lu bee jindi duuma.*
 1.SG my mother GEN book-DET-PL all carry down
 ‘I carried down all of my mother’s books.’

Several of the participants explained the difference in terms of *kiliŋ kiliŋ* ‘one one’. Here is what one of the participants said verbatim about the two sentences above, (18a) referring to the example (16) with the X-woo-X construction, and (18b) providing judgement about example (17).

- (18) a. *Ñiŋ fraaz foloo, i ye i kiliŋ kiliŋ jindi le,*
 this sentence first 2P.SG PRED 3P.PL one one carry-down PERF
 ‘This one you carried them down one by one, ...’
 b. *ñiŋ do, i ye i bee le jindi ñoŋ na.*
 this some, 2P.SG PRED 3P.PL all FOC carry_down together OBL
 ‘... this other one, you carried them down all together.’

These remarks further support the idea that ‘X’ in X-woo-X constructions is the share in an event-key distributive relation.

Overall, there is good evidence that X-woo-X signals the existence of multiple subevents. These findings support a view on which ‘X’ in an X-woo-X construction picks out the share in

an event-key distributive relation. If so, then Gil’s conjecture is upheld in Mandinka; this nominal reduplication construction marks the share.

But if that is the case, then why is it translated as ‘every’? In other words, why does the construction communicate exhaustivity with respect to the X’s? In the next section, we will give evidence that exhaustivity is indeed part of the meaning of X-woo-X, and then develop a hybrid analysis on which X is simultaneously share and key.

3 Exhaustivity effect

To confirm that the X-woo-X construction conveys exhaustivity, we asked for truth value judgments relative to exhaustive and non-exhaustive displays, with X-woo-X in various grammatical positions. Our methodology was inspired by the work of Bosnić et al. (2022) on Serbian *po*, who collected truth value judgments on that distributivity marker relative to exhaustive and non-exhaustive displays.

For this study, we distributed 6 different surveys evenly to 12 native speakers of Mandinka (the Phase II participants described above). Each survey contained two questions, one with an exhaustive display, and one with a non-exhaustive display. Both questions were about a sentence with X-woo-X in the same grammatical position (subject, object, or both). Order of exhaustive vs. non-exhaustive was counterbalanced, so that half of the participants saw the exhaustive display first, and then saw the non-exhaustive display, and the other half saw the displays in the opposite order. The study was thus a $3 \times 2 \times 2$ design, with grammatical position and order as between-participants factors, and exhaustive vs. non-exhaustive as a within-participants factor.

Exhaustivity in subject position. To test exhaustivity in subject position, we used the display in Figure 4, where every town has a doctor, but not every town has a nurse.



Figure 4: Display for testing exhaustivity in subject position (multiple towns, all having a doctor).

Participants were asked to judge the truth of the following sentences.

- (19) **Saatee-woo-saatee** ye jarar¹laa soto le.
 town-DIST-town PRED worker have PERF
 ‘Each town has a doctor.’
- (20) **Saatee-woo-saatee** ye karandir¹laa soto le.
 town-DIST-town PRED teacher have PERF
 ‘Each town has a teacher.’

Participants were given three possible options as responses to choose from: (i) *Tonya loŋ* ‘True’; (ii) *Tonya nteŋ* ‘Not true’; and (iii) *A manke tonya ti, a manke fanya ti* ‘Not true, not a lie’.

Relative to the display in Figure 4, 4/4 participants said that (19) was true, because indeed, every town has a doctor in the display. If we change the noun from ‘doctor’ to ‘teacher’, as in (20) then the sentence becomes false, because not every town has a teacher.

Exhaustivity in object position. To test exhaustivity in object position, we used the two displays shown in Figure 5. Relative to these two displays, participants were asked to judge the truth of the following sentence:

- (21) Saate-e ye **dookuulaa-woo-dookulaa** soto le
 town-DET PRED worker-DIST-worker have PERF
 ‘The town has every (kind of) worker.’

Relative to the exhaustive display in Figure 5 (left), (21) was judged true by 4/4 participants, as the town does indeed have every type of worker. The same sentence is unanimously judged as false in the non-exhaustive display (right), where the town does not have every type of worker.

Exhaustivity in both subject and object position. Finally, we collected judgments on a sentence with X-woo-X in both subject and object positions:

- (22) **Saatee-woo-saatee** ye **dookuulaa-woo-dookulaa** soto le
 town-DIST-town PRED worker-DIST-worker have PERF
 ‘Every town has every (kind of) worker.’

We asked for truth value judgments on (22) relative to the two displays shown in Figure 6.

Example (22) was unanimously (4/4) judged true relative to the exhaustive display in Figure 6 because, indeed, each town has all the different types of workers. In the non-exhaustive display, where not every town has every type of worker, the same sentence is judged false.

Summary and discussion. The findings from the exhaustivity study are very clear and simple: When the display is exhaustive, the sentence is true; with a non-exhaustive display, the sentence is false. Hence X-woo-X is interpreted exhaustively with respect to X, at least in argument position.³ In this respect, the ‘X’ in ‘X-woo-X’ behaves like the key in a distributive relation than the share.

³There are adverbial uses of X-woo-X that do not appear to be exhaustive, as in *luŋ-oo-luŋ* ‘every day’ or *waati-woo-waati* ‘every time’. Thanks to Jakob Mache for raising this point.

4 Analysis

We have seen that X-woo-X behaves partly like a share marker and partly like a key marker. In light of this, we propose a hybrid analysis. To build up to that, let us begin with a treatment of X-woo-X as a share-marker in an event-semantic framework, and let us concentrate on the simple example in (23).

- (23) **Moo-woo-moo** naata le.
 person-DIST-person come PERF
 ‘Everybody came.’

On our share-marker analysis, this sentence describes an event that can be divided into subevents whose agent is a person, which are coming events. Formally, this can be represented as in (24): it’s a property that holds of event e if e is a sum of person-coming events.

- (24) $\lambda e. e \in * \lambda e' [\text{p}(\text{ag}(e')) \wedge \text{come}(e')]$

Based on this example, we can extrapolate a lexical entry for *-woo-* on which it takes a property P and a thematic role θ (such as ‘agent’) and an event description V (such as the property of being a ‘coming’ event) and gives back a property that holds of an event e if it is the sum of V -ing events whose θ -participant has property P .

- (25) $-woo- \rightsquigarrow \lambda P \lambda \theta \lambda V \lambda e. e \in * \lambda e' [P(\theta(e')) \wedge V(e')]$ [first attempt]

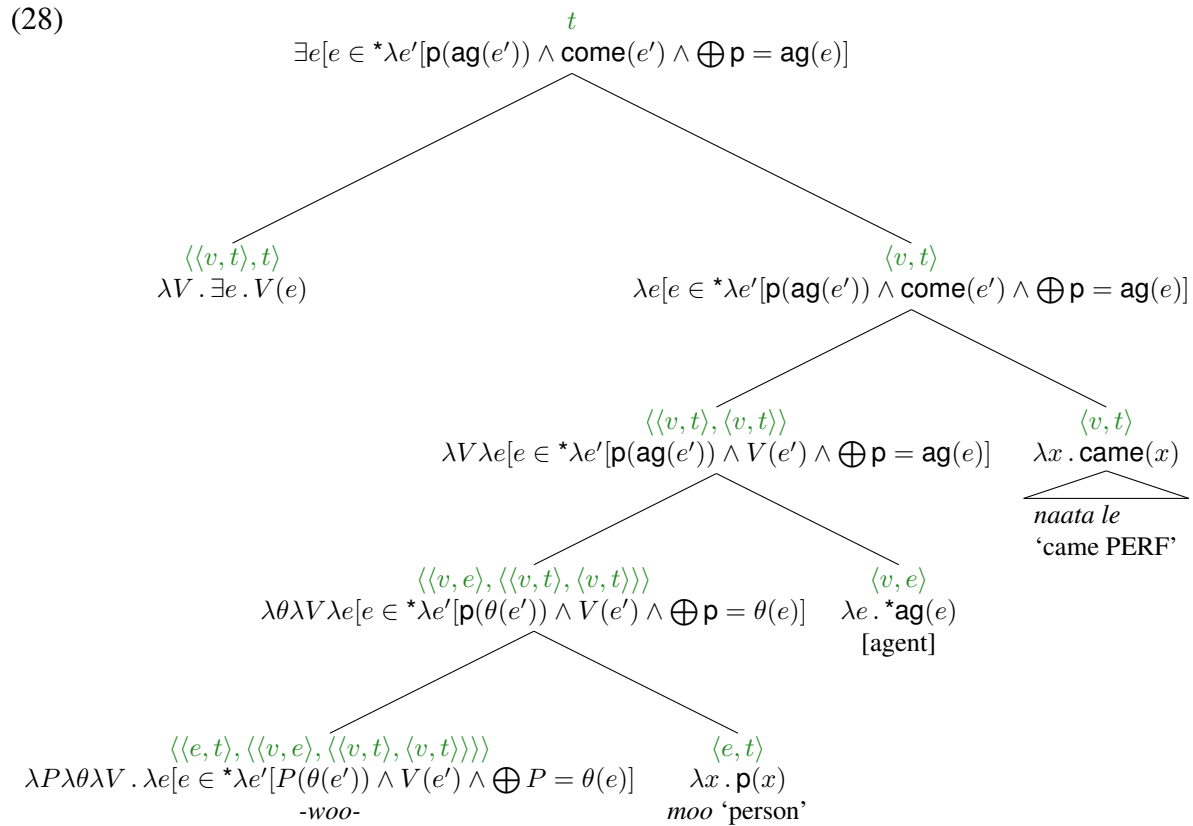
This analysis makes ‘X’ the share in an event-key distributive relation, and thus captures the one-by-one effects. But so far we have not introduced anything into the analysis that would guarantee exhaustivity. To do that, let us add the requirement that, for example (23), the agent of e is the sum of all the people. The event described in (23) will be an event that is a sum of person-coming events whose agent is the sum of all the people:

- (26) $\lambda e [e \in * \lambda e' [\text{p}(\text{ag}(e')) \wedge \text{come}(e')] \wedge \oplus \text{p} = \text{ag}(e)]$

More generally, *-woo-* will require that the θ -participant of the macro-event e is the sum of all the P s. We incorporate that into our lexical entry for *-woo-* by saying that the sum of the P s is the θ -participant of e .

- (27) $-woo- \rightsquigarrow \lambda P \lambda \theta \lambda V \lambda e [e \in * \lambda e' [P(\theta(e')) \wedge V(e')] \wedge \oplus P = \theta(e)]$ [final attempt]

Compositionally, the derivation proceeds as in (28): *-woo-* combines first with the noun *moo* ‘person’, then with the agent theta role, then with the verbal predicate. At the top, existential closure applies to form an expression of type t .



To summarize: We propose that *-woo-* is a hybrid between a share marker and a key marker. This analysis captures both the one-by-one effect and the exhaustivity property. Insofar as our analysis makes X the share in an event-key distributive relation, we capture the one-by-one effect; but the analysis also encodes universal quantification over the X's, and in that respect X is like the key in a distributive relation.

This analysis implies that the Mandinka X-woo-X construction is an instance of ‘simultaneous distributivity’ as Henderson (2019) calls it, since it imposes constraints on both a nominal argument and an event key. Henderson cites another example of this from Mellesmoen’s (2018) work on Comox Sliammon. As Henderson points out, the existence of this phenomenon “degrades the key-share relationship” (Henderson, 2019, 14).

It turns out that the proposed lexical entry is more or less identical to Champollion’s (2016) analysis of determiner *each* and Kuhn & Aristodemo’s (2017) analysis of EACH in French Sign Language. Unlike *every*, English *each* requires different subevents (Tunstall, 1998; Brasoveanu & Dotlačil, 2015; Thomas & Sudo, 2016). English *each* has been observed to be subject to an event differentiation requirement, which can be brought out using the continuation *...but not individually*:

(29) Jake photographed (every / #each) student in the class, but not individually.

We found a similar effect in Mandinka with X-woo-X:

(30) # Jake ye **dindiŋ-oo-dindiŋ** fotoo le, bari a **maŋ** a ke **kiliŋ kiliŋ**
 Jake PRED kid-DIST-kid photog. PERF, but 3SG NEG 3SG DO one one
 ‘Jake photographed each kid but not one by one.’

Furthermore, unlike English *every*, English *each* is unacceptable with *almost* (Farkas, 1997).

(31) Almost every / *each student left the room.

Our *each*-like treatment of X-woo-X predicts that it should be unacceptable with a translational equivalent of *almost*. That predication is borne out; we get similar effect with Mandinka X-woo-X:

(32) *Fode ye **pereske** siise-e **kili-woo-kili** samba le
 Fode PRED almost chicken-DET egg-DIST-egg carry PERF
 ‘*Fode carried almost each egg.’

These parallels suggest that our analysis is on the right track.

5 Conclusion and outlook

We have argued for and presented a hybrid analysis of the Mandinka X-woo-X construction, on which it simultaneously marks the key and the share in a distributive relation. Supporting evidence for this analysis has come from the one-by-one and exhaustivity effects that we have found, along with further parallels between X-woo-X and determiner *each* suggesting that both involve a subdivision of the event into subevents that uniquely correspond to instances of the associated noun.

One fact that remains unexplained is the ‘different kinds effect’ that we found in our investigation of the one-by-one effect. Recall from Section 2 that X-woo-X was judged acceptable in the all-at-once scenario as long as there were different kinds of eggs. We will not offer a full account of this observation here, but our tentative suggestion is that perhaps X-woo-X depends on an ordering on the set of X’s, and that X-woo-X constructions involve a progression along that ordering (cf. Henderson 2013 on English *X by X* constructions). To complete the explanation, it would be necessary to assume further that types can be ordered, while individual eggs are not ordered as easily.

We leave it to future work to flesh out this idea, along with a number of other things to investigate in the future. We mentioned in footnote 3 that adverbial uses of X-woo-X do not appear to carry an exhaustivity effect. Scope is another issue to investigate; X-woo-X appears to take wide scope relative to negation obligatorily. For example, the following sentence only has a $\forall > \neg$ reading, paraphrasable with *no* (as in *saw no animals*):

(33) Jato-o **maŋ daafɛŋ-oo-daafɛŋ** je bii.
 lion-DET NEG animal-DIST-animal see today
 ‘The lion saw no animals today.’

Another direction for future work is to look at similar constructions in other languages. There are other Mande language that have an X-woo-X construction. The following is an example from Dan-Gɛɛtaa (South Mande) (Vydrin, 2017):

(34) **Bɛ ɔ́ ɔ́ ɔ́** ý dū, ā dō à ɓà-’. (Dan-Gɛɛtaa)
 human DIST human who comes I go 3SG beat-INF
 ‘Whoever comes, I’ll beat him/her.’

X-woo-X exists in Jahanke and Bambara too (personal observation by first author). These languages are part of the Mande language family. Wolof, a non-related language but spoken in the area, is reported to have this construction as well (Tamba et al., 2012). Consider the following example.

- (35) a. **Góór-óó-góór** ma giskó [Wolof]
 man-oo-man 1SG see-3SG
 ‘I saw every single man’
 b. Dem-na-a **kër-óó-kër**
 Go-FIN-1SG house-oo-house
 ‘I went to every single house.’

Gilman (1986, 40) mentions a number of apparently related cases of nominal reduplication in African(-diasporic) languages, including *peni peni* ‘a penny each’ in Engenni, Kikongo *kimosi kimosi* ‘one by one’, *wan wan* ‘one by one, one each’ in Cameroonian Creole, *dosu dosu* ‘two each’ in Príncipe Creole and *dé dé* ‘two by two’ in Haitian Creole.⁴

Despite all the work that remains to be done, we have made some progress. Gil (1995) asked “whether there exist *bona fide* instances of reduplication with the interpretation of distributive-key universal quantifier”, and in some sense, we have answered this question in the affirmative. That is, nominal reduplication in Mandinka does have the interpretation of distributive-key universal quantifier, although it is simultaneously a share-marker. On the other hand, Gil’s conjecture remains a possibility: It could be that whenever a reduplication construction serves as a distributivity marker, it marks the share in a distributive relation. It remains to be seen whether there exist *bona fide* counterexamples to this generalization.

References

- Balusu, Rahul. 2006. Distributive reduplication in Telugu. In *North East Linguistic Society (NELS)*, vol. 36, 39–53.
- Bosnić, Ana, Hamida Demirdache & Jennifer Spenader. 2022. Exhaustivity and homogeneity effects with distributive-share markers: Experimental evidence from serbian po. *Syntax* 25(1). 1–38.
- Brasoveanu, Adrian & Jakub Dotlačil. 2015. Strategies for scope taking. *Natural Language Semantics* 23. 1–19.
- Champollion, Lucas. 2016. Overt distributivity in algebraic event semantics. *Semantics & Pragmatics* 9(16). 1–65. doi:10.3765/sp.9.16.
- Choe, Jae-Woong. 1987. *Anti-quantifiers and a theory of distributivity*: University of Massachusetts at Amherst dissertation.
- Farkas, Donka. 1997. Dependent indefinites. In Francis Corblin, Danièle Godard & Jean-Marie Marandin (eds.), *Empirical issues in formal syntax and semantics*, 243–268. New York: Peter Lang.

⁴Thanks to Jakob Mache for pointing us to this source.

- Gil, David. 1995. Universal quantifiers and distributivity. In Angelika Kratzer Emmon Bach, Eloise Jelinek (ed.), *Quantification in natural language*, 321–362. Dordrecht: Kluwer.
- Gil, David. 2013. Conjunctions and universal quantifiers. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*, Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/56>.
- Gilman, Charles. 1986. African areal characteristics: Sprachbund, not substrate? *Journal of Pidgin and Creole Languages* 33–50. doi:10.1075/jpcl.1.1.04gil.
- Henderson, Robert. 2013. Quantizing scalar change. In Todd Snider (ed.), *Proceedings of Semantics and Linguistic Theory (SALT 23)*, 473–492. Linguistic Society of America.
- Henderson, Robert. 2019. Pluractionality and distributivity. Ms., University of Arizona, indicated as forthcoming in *Handbook of North American Languages*.
- Kuhn, Jeremy & Valentina Aristodemo. 2017. Pluractionality, iconicity, and scope in French Sign Language. *Semantics & Pragmatics* 10(6). 1–49.
- Mellesmoen, Gloria. 2018. A one (morpheme) by one (morpheme) approach to paʔapyaʔ:-vʔ- as a temporal pluractional infix in comox-sliammon. In *Proceedings of the 53rd international conference on salish and neighbouring languages*, ed. marianne huijsmans, roger lo, oksana tkachman, and daniel reisinger, 143–60.
- Moravcsik, Edith. 1978. Reduplicative constructions. In J. H. Greenberg (ed.), *Universals of human language*, Stanford: Stanford University Press.
- Tamba, Khady, Harold Torrence & Malte Zimmermann. 2012. Wolof quantifiers. In Edward Keenan & Denis Papern (eds.), *Handbook of quantifiers in natural language* Studies in Linguistics and Philosophy, 891–939. Springer Netherlands.
- The Joshua Project. 2022. Mandinka in Senegal. joshuaproject.net/people_groups/13491/SG.
- Thomas, Guillaume & Yasutada Sudo. 2016. Cumulative readings of *each*. Presentation at the Workshop on (Co-)Distributivity, Paris, February 11-12, 2016. <https://www.ucl.ac.uk/~ucjtdo/pdf/eachParis2016.pdf>.
- Tunstall, Susanne Lynn. 1998. *The interpretation of quantifiers: Semantics & processing*: University of Massachusetts at Amherst dissertation.
- Vydrin, Valentin. 2017. Quantifiers in Dan-Gwɛetaa (South Mande). In Edward Keenan & Denis Paperno (eds.), *Handbook of quantifiers in natural language: Volume ii*, Berlin: Springer.

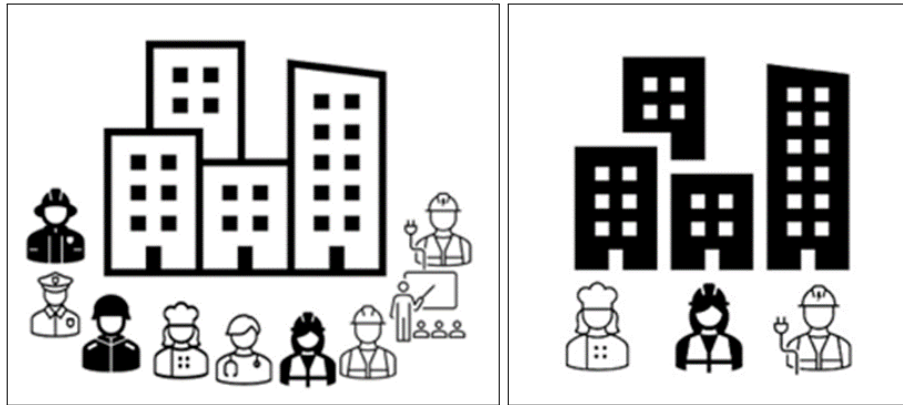


Figure 5: Displays for testing exhaustivity in object position. Left: Exhaustive display. Right: Non-exhaustive display. Both images were shown individually on a slide, accompanied by the array of workers shown above.

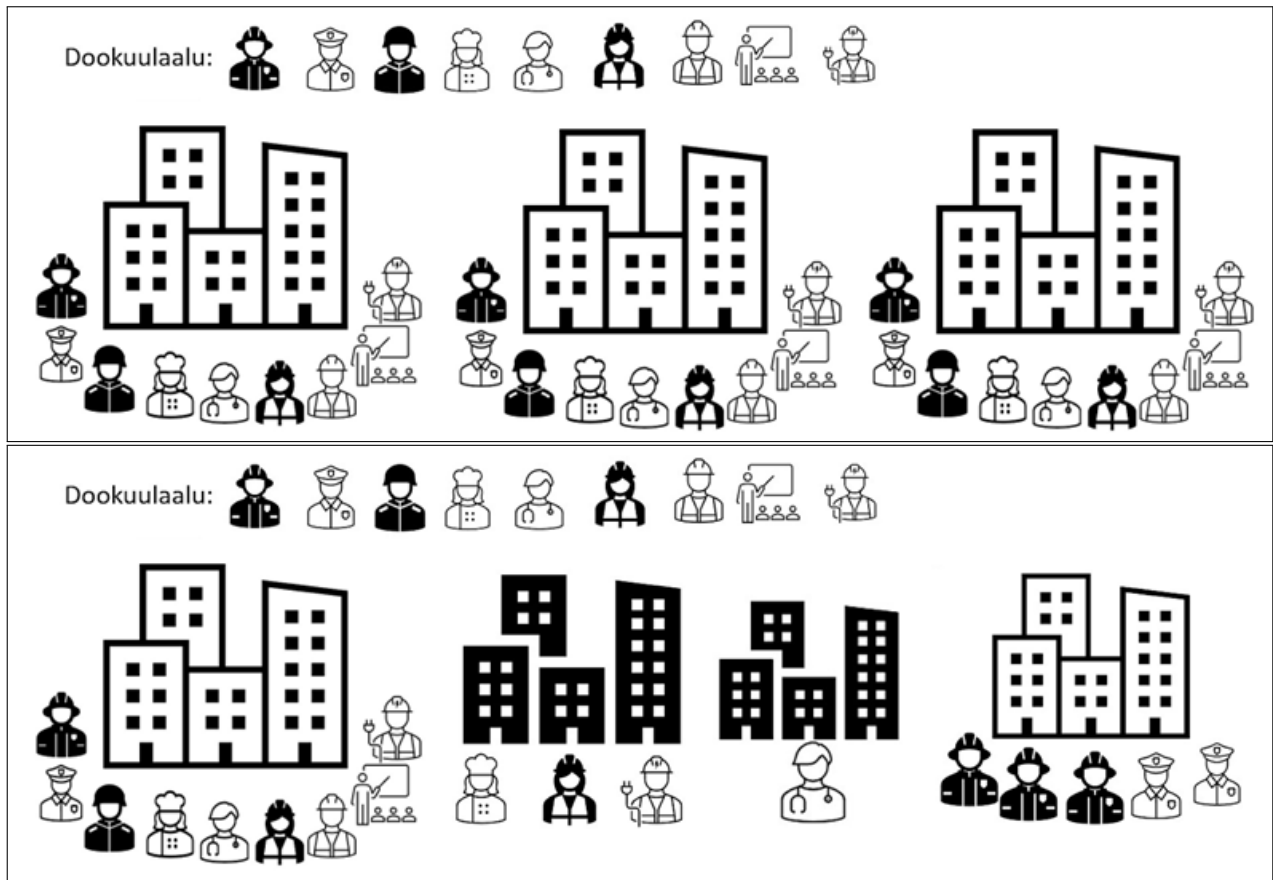


Figure 6: Displays for X-woo-X in both subject and object position. Above: Exhaustive. Below: Non-exhaustive.

Indexed definiteness without demonstratives in Guébie¹

Badiba Olivier Agodio — Guébie Community Member

Peter Jenks — UC Berkeley

Hannah Sande — UC Berkeley

Malte Zimmermann — Universität Potsdam

Abstract. In this article we describe the semantic distribution of the Guébie definite enclitic =*a*, which occurs in a subset of unique definite contexts as well as in anaphoric contexts. We further show that Guébie completely lacks an exophoric demonstrative, the only known counterexample to the proposed universal that every language has demonstratives. We analyze =*a* as an indexed unique definite, a novel category of definiteness, and bare nouns are inherently indefinite. The Guébie facts suggest that while demonstratives are not universal, indexed definiteness is.

1 Introduction

In this paper we provide a description and first analysis of definiteness marking in Guébie (Kru). The central novel observations that we report on are as follows:

The first finding we report on is that there are two morphosyntactic forms which occur in definite contexts in Guébie, i.e. in contexts involving uniqueness and/or anaphoricity. The first form is an enclitic definite article =*a*, which is used in many anaphoric contexts, but which differs in meaning and distribution from standard familiarity definite articles discussed in Schwarz (2009). The second form is a bare NP, which shows up in some unique definite contexts, though not in others. We will analyse such bare NPs as underlyingly indefinite, i.e. non-indexed and non-unique, despite their occurrence in some uniqueness contexts.

Our second central finding is that Guébie lacks an exophoric demonstrative entirely, thereby constituting the first known exception to the proposed absolute universal: “All languages have demonstratives. . .” (Diessel, 1999, p.1).

Crucial to the finding that =*a* is not a demonstrative is the observation that it occurs in (some) uniqueness contexts where demonstratives are infelicitous, and that it is infelicitous in contexts licensing (pointing) demonstratives. In response to this, we will propose a new analysis that features a novel type of definite determiner. We will argue that Guébie =*a* is felicitous in contexts that independently satisfy the following two conditions: (i.) direct (anaphoric or exophoric) reference to a contextually salient individual, and (ii.) situational uniqueness. This is formalised in (1):

$$(1) \quad \llbracket = a_y \rrbracket^g = \lambda s. \lambda P_{\langle s, et \rangle}: \exists!x [P(x)(s)] \wedge \iota x [P(x)(s)] = g(y). \iota x [P(x)(s)]$$

According to (1), the Guébie definite marker =*a* is an *indexed unique definite*, combining components of unique and indexed DEFs. It comes with two presuppositions: (i.) that there be just

¹We would like to thank all our Guébie consultants for their patience and insightful comments on their language, as well as Katherine Russell for her help in eliciting data with additional speakers. Thanks also to the audiences of the 2nd Definiteness Network Meeting in Potsdam and TripleA10 in Potsdam for fruitful discussion. This work received support from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project ID 317633480 – SFB 1287, Project C02, and NSF CAREER Award #2236768.

one individual with the NP-property P in the evaluation situation (*UNIQueness*); and (ii.) that this unique individual be identical to some contextually provided individual $g(y)$ (*INDEXicality*). The two presuppositions combined make indexed $=a$ semantically stronger than plain uniqueness definites, which denote a unique NP-instance irrespective of indexicality, on the one hand, and familiarity definites à la Schwarz (2009), which denote the unique prementioned NP-instance, on the other. These stronger semantics correlate with a more constrained distribution of $=a$ compared to plain uniqueness or anaphoric DEFs. In contrast, bare NPs are always indefinite and used when neither of the context conditions *UNIQU* or *INDEX* is met. Finally, the different behavior of Guébie $=a$ in situation-based covariation and donkey sentences may constitute a novel empirical argument for a dynamic DRT-type analysis of donkey sentences à la Kamp (1981).

The article is structured as follows. Section 2 provides some background on unique and anaphoric definite markers. Section 3 introduces the relevant empirical data on how Guébie divides the semantic task of expressing uniqueness and anaphoricity between bare NPs and definite NPs with $=a$. We also show that $=a$ is not a demonstrative marker, and that there are no other demonstratives in the language. Section 4 presents the formal analysis of $=a$ and bare NPs in Guébie. Section 5 concludes.

2 Some Background: Anaphoric vs unique definites

The literature offers ample discussion of two subtypes of definites, viz. *uniqueness* definites and *anaphoric* definites. For instance, overt definite markers have been argued to code anaphoricity, rather than uniqueness in the West African languages Hausa (e.g., Newman, 2000, Zimmermann, 2008) and Akan (Arkoh and Matthewson, 2013, Owusu, 2022; pace Bombi, 2018). There are two analyses of anaphoric (aka familiar, strong) definites, here exemplified for the Akan DEF-marker *nó*. In both theories, there are special mechanisms for weakening or avoiding a uniqueness presupposition. The first account was originally proposed in Schwarz (2009) and takes anaphoric definites to presuppose uniqueness relative to some index (Arkoh and Matthewson, 2013):

$$(2) \quad \llbracket n\acute{o}_y \rrbracket^g = \lambda s. \lambda P_{\langle s, et \rangle}: \exists! x [P(x)(s) \wedge x = g(y)]. \iota x [P(x)(s) \wedge x = g(y)]$$

According to (2), *nó*-NPs can be used if the context is such that there is exactly one individual with property P that is identical to some established discourse referent $g(y)$. Such *nó*-NPs then refer to this unique indexed individual, thereby weakening the uniqueness requirement.

Owusu (2022, p. 48), in turn, takes up an idea for demonstratives by Dayal and Jiang (2022) and treats the Akan DEF-marker *nó* as an adnominal modifier encoding familiarity and non-uniqueness of the NP-restriction:

$$(3) \quad \llbracket n_y \rrbracket^g = \lambda P_{\langle s, et \rangle}. \lambda x. \lambda s: x = g(y) \wedge \exists s' [s \leq s' \wedge |\{x: P(x)(s')\}| \geq 1]. P(x)(s)$$

According to (3), *nó* can be used in a context where there is a discourse referent in a situation s , which is itself part of an extended super-situation s' containing more than one individual with property P . This directly rules out the use of *nó* in plain uniqueness contexts.²

²The formulation in (3) poses an immediate problem: Without additional restrictions on the super-situation it seems to over-generate the distribution of *nó*: Since there is almost always a super-situation containing more than

Both of these analyses avoid making the uniqueness component of the anaphoric definite too strong, as definites and demonstratives often occur in contexts where absolute uniqueness is not met. In contrast, we will show below that Guébie $=a$ only occurs in contexts where uniqueness is met in an absolute sense, relative to some situation s , calling for the analysis in (1).

3 The expression of definiteness in Guébie

Guébie (Eastern Kru) is spoken by about 7000 speakers in the Gagnoa prefecture of Côte d’Ivoire; see Sande (2020) for a language snapshot and references. The basic word order is SVO and SAuxOV, and it has a 4-height tone system. All of the primary data come from face-to-face and online elicitations with our co-author, Badiba Olivier Agodio, a 40-year-old male native Guébie speaker. The data were subsequently checked with two other male speakers, one older and one younger, who confirmed the original judgments with minimal variation. In addition, we examined several texts to ensure that the elicitation-based generalizations are sound. All of the examples use the IPA for transcription, so, e.g., [j] is a palatal glide and [c] is a voiceless palatal stop. They also use Leipzig glossing. The data are available in the Guébie collection of the California Language Archive (Agodio et al., 2014).

In what follows, we will first look at the realisation of NPs in typical unique definiteness contexts, before turning to the realisation of NPs in typical anaphoric contexts. We will see that Guébie sometimes uses bare NPs and sometimes NPs DEF-marked with $=a$, but that the distribution of bare NPs and DEF-marked NPs is different from that of other languages. Finally, we will report on the infelicity of $=a$ in demonstrative (pointing) contexts, which violate situational uniqueness.

3.1 Unique definiteness contexts: Bare NPs vs. $NP=a$

Following Hawkins (1978), Schwarz (2009), Jenks (2018), i.a., we look at the formal realisation of NPs in four typical unique definiteness contexts: larger situations; immediate situations; situational covariation; and part-whole bridging. In languages that have them, these contexts typically require a plain uniqueness DEF-marker.

With cases of *larger situation uniqueness* (e.g. *sun, moon, president*), Guébie speakers prefer bare NPs, cf. (4-a), as is the case in many other bare noun languages, such as, e.g., Mandarin (Jenks, 2018) or Babanki (Akumbu and Jenks, 2023). However, such larger situation unique definite NPs can also be marked by DEF $=a$, as long as they (i) have a salient exophoric discourse referent (Barlew, 2014), or (ii) are anaphoric, cf. (4-b):

- (4) a. jiro-je^{2.3.1} pɔ¹ (only unique: bare)
 sun-SG shine
 ‘The sun is shining.’
 b. jiro-je=**a**^{2.3.1.1} pɔ¹ (unique & indexed: =a)
 sun-SG=DEF shine
 ‘The sun (that we were waiting for) is shining.’

one individual satisfying the NP-predication, the occurrence of *nó* should be almost vacuously licensed. Given the multitude of suns in the universe, *nó* should then freely occur in (4-a) below, contrary to fact.

Out of the blue, the overt definiteness marker =*a* was dispreferred in (4). The example was first given without an overt definite. When asked if the example with =*a* was okay, the speakers accepted (and repeated) it, but they offered extra context to make the DEF-marker felicitous. The addition in parentheses in the paraphrase of (4-b) suggests that the unique sun is prementioned or otherwise indexically accessible in the discourse. As there is an established discourse referent, =*a* is felicitous. The same holds for examples with *moon*, not shown here. The following example in (5) tests specifically for the felicity of =*a* in a context aimed at making the sun not perceptually accessible, and hence non-indexed. The speaker's comment is telling, as it shows that the invisible sun must be accommodated as a discourse referent, here by its effect on the room's temperature. Again, this shows that =*a* is indexed and in need of an accessible discourse referent to refer to.

- (5) Context: We're in a dark room with no windows and no lights and someone says:
 jiro-je=**a**^{2.3.1.1} tɛ(-a)^{3.2} me³
 sun=DEF be.strong-PST PART
 'The sun is/was strong.'
 Speaker's comment: 'Well, in that case it would be really really hot in the room...'

NPs in *immediate situation contexts* show a similar effect. If the intended unique referent is unfamiliar or new to the conversation, the bare noun is used. This is shown in (6-a) with the NP *dirɛtɛ* 'headmaster', for which the bare NP was offered as the original translation. Adding the DEF-marker =*a* is only felicitous to distinguish between different schools or directors of different schools. In this case, speakers seem to be accommodating an anaphoric interpretation of =*a*.³

- (6) Context: A father is visiting a new school for his son, and someone says:
 a. dirɛtɛ(=**a**)^{4.1.1.1} ɔ³ lɔpɛ^{2.1} ɓɔ³
 director.FR=DEF 3SG.NOM speak.IPFV finish
 'The director will speak to you after (whatever he's doing now).'
 b. sukulu-masi^{1.1.3.4.1} ɔ³ ji³ mɛ-sale=da^{4.2.3.2} kɔ³ ji³
 school-master 3SG.NOM FUT PART-talk=place.NMLZ to come
 'the/a teacher will come to talk to you.'

Crucially, (6-a) differs from (6-b) with *sukulu-masi* 'teacher' in that the latter is infelicitous with =*a* if there is more than one teacher, thereby violating uniqueness, or if there has been no talk about a specific teacher, thereby violating indexicality.

At the same time, the fact that =*a* is often available in uniqueness contexts with additional implications about contextual relevance or salience is surprising from the perspective of anaphoric definites in, e.g., Mandarin or Thai, where anaphoric demonstratives are awkward with unique referents (Jenks, 2015; Jenks, 2018). This suggests that =*a* is not a demonstrative as in Dayal and Jiang (2022), a point to which we will return in sub-section 3.3. The contrast extends to instances of exophoric reference in immediate contexts with a salient unique discourse referent (cf. Barlew, 2014), which require =*a* for felicity.

- (7) Context: Djatchi goes to Gnadja's house, and, as he comes in, he notices a dog lying in the

³Alternatively, this might be an instance of part-whole-bridging from 'school' to 'director'; see main text below.

corner sleeping. He says:
 goji=#(a)^{3.1.1} ɔ³ nanɛ^{3.3}
 dog=DEF 3SG.NOM be.good
 ‘The dog is nice.’

In (7), there is no larger situation in which the dog is not unique. The context supports only the existence of a single unique dog, who is also present in the situation. Such contexts typically support exophoric uses of demonstratives (regardless of whether the referent is contextually unique), but they do not support dedicated anaphoric demonstratives, cf., e.g., Akumbu and Jenks (2023).

The remaining two typical unique definite contexts show that =*a* requires indexed AND unique referents. Firstly, in part-whole bridging, =*a* is preferred because the bridging context makes the situationally unique referent anaphorically accessible. Example (8), from a narrative, illustrates:

- (8) Context: Description of how to build a house. Previous sentences describe pouring concrete for the floor and lower walls. There has been no mention of the ‘top’.

ɔ³ ka³ jɛ³ nɛ² wa³ nɔ² wɔli=#a^{3.1.1}
 3SG.NOM IRR dry REL 3PL.NOM do.IPFV top=DEF
 ‘When it’s dry, they do the top.’

Source: House-building text, spoken by Badiba Olivier Agodio on 8/11/2016

The fact that =*a* is used in such contexts licensing uniqueness definites in other languages (e.g., bare nouns, weak articles in German), strongly suggests that it is compatible with uniqueness, and that it does not require previous mention in the literal sense. Nevertheless, it refers to a contextually-accessible discourse referent in such contexts, presumably via accommodation.

Finally, contexts with situation-based covariation trigger the use of unique definite NPs in many languages (Schwarz, 2009). In Guébie, though, we find that both =*a* and bare (singular) NPs are infelicitous in such contexts, which require a bare plural NP instead. The consultant’s comment suggests that =*a* is impossible because there is no specific **unique** discourse referent for NP=*a* to refer to (and that situation-binding of =*a*’s individual index variable is impossible).

- (9) Context: In every church we visited, we spoke with the priest.
 lagɔ^{3.1} bitə^{2.2} mɛ³ tu⁴¹ la² mɔ² e⁴ jeralɪ^{3.2.2} anɛ^{2.2} e=#a^{4.1} mɔna^{2.2}
 god house in all of place 1SG.NOM visit REL 1SG.NOM=PST ?
 (lagɔ-ŋw-a^{3.12}) **gali**^{3.1} lɔpɛ^{3.1}
 god-AGT-PL elder.PL speak

‘In every church I visited I spoke with the priests/elders.’

Comment: ‘=*a* is okay on ‘elder’ only if you visited one church and spoke to one elder/priest there, and it’s someone you know about or have been talking about already.’

Summarising our findings for the realisation of NPs in unique definite contexts in Table 1, we see that the pattern of Guébie DEF-marking clearly differs from the Mandarin one. In particular, the distribution of the Guébie DEF-marker =*a* extends to several unique definite contexts. In contrast, Mandarin only allows bare nouns in these contexts, never demonstratives. Secondly, the unavailability of bare SG NPs in both part-whole bridging and situation-based covariation contexts suggests that bare SG NPs in Guébie differ semantically from their Mandarin counterparts, which

are commonly assumed to type-shift via iota (Dayal, 2004; Jiang, 2012; Jiang, 2020; Jenks, 2018).

Context	Guébie	Mandarin
Larger situation	Bare NP, (=a)	Bare NP
Immediate situation	Bare NP, (=a)	Bare NP
Situation binding	Bare PL NP	Bare NP
Part-Whole	=a	Bare NP

Table 1: NPs in uniqueness definite contexts

3.2 Anaphoric contexts: $NP=a$

In Guébie, the use of $=a$ is obligatory in most anaphoric definite environments. To begin with narrative sequences, which constitute the canonical uses of anaphoric definites, the DEF-marker $=a$ is obligatory, as shown in (10-ab). Crucially, $=a$ is strange at the first mention of ‘president’ in (10-a), but it is strongly preferred once the discourse referent is established in (10-b). This highlights once more the indexed meaning component. Moreover, the felicity of $=a$ on the uniquely referring singleton-NP ‘president’ shows once more that the DEF-marker is compatible with contexts where uniqueness presuppositions are satisfied.

- (10) Context: There was a ceremony in the village yesterday, and I say:
- a. kuβə^{2.31} ane^{4.2} cifi-ɲɔ^{4.1.2} ɔ³ la=a^{31.2} anε-du^{2.3.3} galɪa^{2.3.1}
 yesterday REL chief-AGT 3SG.NOM call.PFV=PST 1PL.POSS-village great
 la² ka-ma-nɪ=ε-ɲɔ^{2.2.2.2.2} (#=a)
 ASSOC have-become-APPL=3SG.ACC-AGT=DEF
 ‘the village chief invited the president to the ceremony.’
- b. ka-ma-nɪ=ε-ɲɔ=a^{2.2.2.2.2} ɔ³ ko² mε³
 have-become-APPL=3SG.ACC-AGT=DEF 3SG.NOM be.LOC DEM.PRO.PLACE
 anε-truli^{2.3.3.1} la² da²
 1PL.POSS-play ASSOC place
 ‘The president attended the ceremony.’

Importantly, anaphoric reference to the referent of uniquely-denoting NPs such as ‘sun’ also requires $=a$, (11-b). Again, such uses are deviant or at least pragmatically odd with demonstratives.

- (11) a. e⁴ ni⁴ jiro^{3.3} bala^{2.2} εja^{3.1} coje^{3.1} ɪ³ ʝe² pε⁴
 1SG.NOM see.PFV sun diminish.PFV and moon 3PL.NOM while sleep.PFV
 jaanε^{2.3.1}
 today
 ‘I saw the sun and moon go down today.’
- b. jiro-ji#(=a)=a^{2.3.1.1.1} jalia^{2.2.2} titi^{4.4}
 sun-SG=DEF=PST be.red very
 ‘The sun was very red.’

In donkey anaphora, which typically license anaphoric definites over unique definites, =*a* is obligatory, at least with animate referents as in (12):⁴

- (12) ɲɔkpɔ^{3.1} ɲɔkpɔ^{3.1} k^wala^{2.2} goji^{3.1} ne² li-ɔ^{2.2} goji=**a**^{3.1.1}
 person person take.care dog REL eat.IPFV-CAUS dog=DEF
 ‘Every person who has a dog feeds the dog.’

The diagnostic of donkey anaphora is a bit problematic when applied to Guébie though. This is because bare NPs turn out to be possible in such contexts, too, as illustrated in (13). We propose that (13) is felicitous on an indefinite construal of the bare NP, which is compatible with the context in (13). We will return to the indefinite status of bare NPs in section 4.3.

- (13) kɔkɔ^{4.4} mɛ³ e⁴ ka³ ɓabɛɛ^{3.3.3} ɛja^{3.1} wɔli^{4.4} dabara^{4.4.4} kɔ³ ɲɔkɔ-ni-ni^{2.3.4.2}
 everyday in 1SG.NOM IRR sheep and goat market at PART-see-APPL
 wɔli(=**a**)^{4.4} e^{1.4} ɲja^{3.1} (=a optional)
 goat=DEF 1SG.NOM buy
 ‘Every time I see a sheep and a goat at the market, it’s **a/the** goat that I buy.’

A final surprise comes with producer-product bridging, which typically licenses anaphoric definites in other languages (cf. Schwarz, 2009). We would therefore expect Guébie =*a* to be licit in these contexts if it were a run-off-the-mill anaphoric DEF-marker. However, it turns out that =*a* is impossible in Guébie in such contexts. When presented with the Guébie sentence in (14), the consultants did not like the sentence and removed the =*a*. When prompted again if the sentence would be okay with =*a* on ‘author’, they said ‘no’.

- (14) ɲaci^{23.1} ɲja³¹ bagɔ^{3.2} la² lilelu^{3.3.1} ɔ³ wa² cɛli-ɲɔ(#=**a**)^{3.3.2.2} ɲa³¹
 Djatchi buy.PFV book GEN new 3SG.NOM like write-AGT=DEF because
 ‘Djatchi bought a new book because he likes the author.’

Our main results from this and the preceding section are summarised in Table 2 on p. 8. Once again, we see that the distribution of Guébie =*a* is different from that of demonstratives in Mandarin. Generally speaking, Guébie =*a* is used in all contexts with explicit reference to a contextually specified, i.e. indexically accessible individual. Crucially, and unlike with anaphoric definites in many other languages, this extends to contexts in which that individual is unique. Moreover, the use of =*a* in donkey anaphora is similar to anaphoric definites in other languages. As for bare SG NPs, these are felicitous in fewer contexts than in Mandarin. In particular, they are unavailable for situation binding and part-whole bridging, suggesting that such NPs do not denote uniqueness definites in Guébie. Finally, the unavailability of =*a* in producer-product bridging is unexpected and requires further investigation.

⁴The data situation is somewhat complicated by the fact that it is also possible, and maybe even preferred, to use the SG pronoun ‘it’ or the plural NP ‘dogs’ in clause-final position. The co-author also definitely prefers ‘Every person feeds his dog(s)’, with a possessive, to any of the above. In addition, there seems to be some speaker variability regarding the choice of =*a* with inanimate (non-human) NPs, such as, e.g., ‘machete’. Whereas the co-author found =*a* infelicitous or optional with such NPs, the other speakers consulted regularly provided and accepted =*a* with animate and inanimate NPs alike.

Context	Guébie	Mandarin
Anaphoric sequences	=a	DEM
Donkey anaphora	=a	DEM
Product-Producer	Bare NP	DEM
Larger situation	Bare NP, (=a)	Bare NP
Immediate situation	Bare NP, (=a)	Bare NP
Situation binding	Bare PL NP	Bare NP
Part-Whole	=a	Bare NP

Table 2: NPs in uniqueness and anaphoric definite contexts

3.3 On the absence of demonstratives in Guébie

We now turn to a typologically notable claim: Guébie seems to have no demonstratives, thereby constituting an apparent exception to the absolute universal by Diessel (1999, p. 1) that “All languages have demonstratives. . . .” For illustration consider the context in (15), which reliably triggers demonstrative markers in languages that have them. In contrast, there appears to be no way to discriminate between type-identical individuals using an exophoric determiner in Guébie. Speakers give lots of alternatives to avoid the demonstratives in the direct translation, including ‘Which dish do you want?’, or they resort to pragmatic resolution. One speaker commented that “You could say ‘Do you want dish=a?’ (holding one dish up). Then if she says ‘no’ you know she wants the other one.” When presented with the sentence ‘Do you want dish=a or dish=a’ in Guébie, speakers did not accept it, but offered (15) with a semantically enriched relativised NP instead:⁵

- (15) Context: Edwige asks you for a dish, but there are two on the table, so you ask ‘Do you want this dish or this dish?’
 ʃɛ^{3.1} ɔ³ kɔ² ʃa²/da³ nɛ² o ʃɛ^{3.1} ɔ³ kɔ² da³ nɛ²
 dish 3SG.NOM be.LOC here/there REL or dish 3SG.NOM be.LOC there REL
 ‘the dish that is here/there or the dish that is there’

Its failure in the consistency test (Löbner, 1985) constitutes further evidence that =a is not a demonstrative, despite its underlying indexed nature. (16) is judged as deviant “unless a single dog is sleeping while making noise and we already know which dog we’re talking about.” This comment nicely brings out the double nature of =a as an indexed unique definite.

- (16) #goji=a^{3.1.1} ŋɔŋɔ^{2.2} nɛ⁴ goji=a^{3.1.1} sa³ kpəli³
 dog=DEF sleep.IPFV and dog=DEF make.IPFV noise
intended: ‘This dog sleeps and that dog makes noise.’

These facts suggest that =a cannot serve the discriminating function of demonstratives. On this, it patterns with the Akan DEF-marker *nó*, which likewise cannot serve the discriminating function

⁵Two other Guébie speakers likewise had trouble translating these examples, but they resorted to a different rescue strategy, namely to borrow demonstratives from another Kru language (*ne* ‘this’ and *kɔ* ‘that’). This pattern resembles the Akan strategy in (17-c) below.

of demonstratives on its own, cf. Bombi (2018, p. 152) and Owusu (2022, 21ff.). This holds even when the DEF-markers are accompanied by a pointing gesture (Ebert et al., 2020), as illustrated in (17-ab). For the discriminating function to succeed, the second instance of *nó* must either be replaced by the distal marker *yi*, cf. (17-c), or else *nó* must be accompanied by the additional demonstrative marker *saa*, cf. (18):

- (17) a. #Abofra **nó** nim adeɛ paa ɛna abofra **nó** abɔn. [Akan]
 child DEF know thing INT CONJ child DEF not.smart
intended: ‘This child is smart, and that child is not smart.’ (Owusu 2022: 22, ex.28)
- b. #Me-pɛ car **nó** n-yɛ car **nó**
 1PL-want car DEF NEG-COP car DEF
 ‘I like that car [pointing at Audi] but not this car [pointing at Renault].’
- c. Me-pɛ car **nó** n-yɛ car **yi**
 1PL-want car DEM NEG-COP car DEM
 ‘I like that car [pointing at Audi] but not this car [pointing at Renault].’
- (18) **Saa** abofra **nó** nim adeɛ paa ɛna **saa** abofra **nó** abɔn. [Akan]
 DEM child DEF know thing INT CONJ DEM child DEF not.smart
 ‘This child is smart, and that child is not smart.’ (Owusu 2022: 22, ex.28)

Summing up, it appears neither Akan *nó* nor Guébie =*a* can (easily) serve as the linguistic support of accompanying pointing gestures on their own, but this is one of the defining properties of demonstratives (Lyons, 1999, Ahn, 2022). One could formally model this by specifying that =*a/no* do not have a semantic argument slot for the pointing argument, cf. Ahn (2022). Alternatively, one could also assume that =*a/no* alone cannot function as semantic dimension shifters that would shift the not-at issue meaning of the pointing gesture to the at-issue content (Ebert et al., 2020). In Akan, this would then be the function of the additional marker *saa*, whereas Guébie would lack a dimension shifter altogether for reasons unknown to us at present.

4 Analysis

This section first presents in subsection 4.1 our analysis of Guébie =*a* as an *indexed unique DEF-marker*. This makes =*a* a novel type of DEF-marker that has not been discussed in the previous literature on DEF-marking in natural language. Subsection 4.2 lays out how the analysis accounts for the data and presents some further correct predictions. Subsection 4.3 briefly looks at the meaning of Guébie bare SG NPs. It argues that such NPs in Guébie are indefinite, unlike their unique-definite bare NP counterparts in better studied languages, such as Mandarin. Subsection 5 then concludes by discussing some more general implications of our analysis for the treatment of (indexed) definites and bare NPs in natural language.

4.1 Guébie =*a* as an indexed uniqueness DEF-marker

In the literature, Schwarz (2009)-style analyses of definites assume that there are two types of definite NPs: (i.) those with uniqueness presuppositions (unique definites) in (19-a); and (ii.)

those with indices, such as, e.g., anaphoric definites, in (19-b).

- (19) a. Unique definites (*the dog*): $\iota x[\text{dog}(x)]$ ‘the unique x s.t. x is a dog’
 b. Indexed definites (*that₁ dog*): $\iota x[\text{dog}(x) \ \& \ x = g(1)]$ ‘the unique x s.t. x is a dog and $x = \text{that}_1$ ’

Jenks and Konate (2022) moreover argue that anaphoric definites are part of a larger category of indexed definites. This larger category includes demonstratives as well as pronominal definites of the ‘we linguists’ variety.⁶

Generally, indexed definites as in (19-b) are used whenever there is a need to point back to a particular individual in the discourse, such as with anaphoric definites, to discriminate between different individuals in a context, such as with exophoric definites, or to bind that index variable, as with donkey anaphora.

Given this, we have seen very good evidence that Guébie $=a$ is an indexed definite as well: (i.) the marker is always used in anaphoric contexts, cf. (10-b); (ii.) it can be used to refer to contextually salient exophoric individuals, as long as they are unique, cf. (7); (iii.) it is always referential, and resists situation-based covarying readings, cf. (9); and (iv.) it is required in donkey sentences, cf. (12).

At the same time, $=a$ differs from typical demonstratives, which likewise have an indexical meaning component, in still including a uniqueness requirement. In particular, it cannot discriminate different individuals in a particular context, cf. (16). And it occurs in a number of contexts where uniqueness holds, even in extended situation or global situations, cf. (4)-(5), and with anaphoric uses of immediate situation definites, cf. (6-a).

As for the formal analysis, recall from section 2 that previous analyses of familiarity-indexed definites explicitly weaken the uniqueness requirement of definites. In Schwarz (2009) and Jenks (2018), uniqueness is evaluated relative to the unique individual denoted by the index, making it essentially vacuous, cf. (2) above. In contrast, Dayal and Jiang (2022) and Owusu (2022) argue that demonstratives/familiar definites explicitly encode anti-uniqueness relative to some larger context of utterance, cf. (3). These treatments cannot account for the distribution of Guébie $=a$, which seems to come with an absolute uniqueness requirement.

We propose that the meaning of $=a$ includes the following two semantic components: (i.) a contentful uniqueness presupposition, which is NOT relativized to the index, and which accounts for the absence of discriminating, demonstrative-like uses; and (ii.) an index, which will force the $=a$ -marked DP to refer, to be anaphoric, or to be dynamically bound. The meaning of $=a$ is repeated in (20) from (1) above. The first conjunct in the presupposition formalises the uniqueness condition, the second one the indexicality condition on the felicitous use of $=a$.

$$(20) \quad \llbracket = a_y \rrbracket^g = \lambda s. \lambda P_{\langle s, et \rangle} : \exists! x [P(x)(s)] \wedge \iota x [P(x)(s)] = g(y). \iota x [P(x)(s)]$$

⁶Guebié has an additional DEF-marking strategy involving a combination of NP and pronoun, which is illustrated in (i), and the distribution and function of which we do not yet understand:

- (i) $e^4 \text{ ji}^3 \text{ [butike}^{2.3.2} \text{ e}^3] \text{ me}^3 \text{ me}^2$
 1SG will store 3SG to go
 ‘I will go to the store.’

According to (20), if there is exactly one x that has property P in situation s , and if this x is also identical to some anaphorically or exophorically accessible discourse referent, then $NP=a$ will pick out this indexed and unique individual.

The intuition behind our analysis seems quite similar to the notion of salience in Barlew (2014), even though in his specific formal implementation of the meaning of the Buli DEF-marker $=te$, uniqueness is again relativized, and thereby weakened, to some salient individual in the context. For Barlew, $sal(x,c)$ in (21) requires that both the speaker and addressee are paying attention to x , where x is a weakly familiar DR in the context at utterance time. Salience and uniqueness condition in (21) are presuppositions for the felicitous utterance of $=te$ in context c .

$$(21) \quad \llbracket =te \rrbracket^c = \lambda P_{\langle e,t \rangle} : \exists ! i \in D_c [P(i) \wedge sal(i,c) \wedge \forall j \in D_c [P(j) \wedge sal(j,c) \rightarrow j = i]]. i$$

Reference to a unique and salient individual would also account for the Guébie cases involving anaphoric or exophoric reference discussed so far. When applied to the part-whole bridging example in (8), this would entail that the top of the concrete is made salient by the overall discourse structure, e.g., in terms of relevant QUDs: Since the preceding discourse chunks in (8) are about the foundation and the lower walls and the concrete, they would make the surface of the concrete salient and hence a suitable referent for $NP=a$. Alternatively, reference to salient individuals may just be a frequently attested subcase of indexed reference, as salience is a sufficient condition for this type of reference. We will leave a more detailed formal comparison of our analysis and a salience-based approach to another occasion.

4.2 Accounting for the data and some further predictions

We have already seen how the analysis accounts for standard instances of anaphoric or exophoric reference when the indexed individual is unique in the evaluation situation. The strongest evidence for the analysis in (20), however, comes from the fact that $=a$ -marking is found on anaphoric singleton-denoting NPs, such as ‘president’ in (10-b) and ‘sun’ in (11), and also with the topic-situationally unique ‘director’ in (6-a). In particular, the licit occurrence of $=a$ on singleton-denoting NPs differentiates Guébie $=a$ from Akan $nó$, and motivates its analysis as an indexed unique definite.

Moreover, as the uniqueness prediction must hold in absolute terms, against the evaluation situation, we make an additional prediction. Anaphoric reference to NPs with (prototypical) non-singleton reference, such as, e.g., ‘man’ should be infelicitous because of a violation of uniqueness. This prediction seems borne out as can be seen from the infelicity of (22) in the following context:

- (22) We were previously talking about a man among other men.
 #jaci^{23.14} le² ɲɔkpɔ(=a)^{3.1.1}
 Djatchi.NEG be person=DEF
intended: ‘Djatchi is not (the) man.’

The consultants judged (22) as weird both with and without $=a$ in this context. Instead, the preferred option is a different sentence like ‘It’s not Djatchi’ that does not provide an opportunity for placing the DEF-marker. Interestingly, (22) is possible with $=a$ in a different context where we are looking for someone, but Djatchi is not this person. We contend that this is a case of

accommodation of an immediate search-situation containing just a single man.

Finally, $=a$ is ruled out in the following context in (23) because of uniqueness and its non-demonstrative nature.

- (23) Context: There are ten identical cats. I point to one and say ‘I want that cat’.
 #e⁴ jira^{2,3} sɛpi=**a**^{2,4,4}
 1 SG.NOM ask cat=DEF
 ‘I want the/that cat’

Again, the consultant’s comments in (23-a) with $=a$ are telling. While not the first option given, the sentence is judged to be okay in a slightly changed situation where there are 10 identical cats and you point to one and say ‘I want this cat’, as long as you are picking up the one cat or making very clear which one it is. This also seems to involve a uniqueness-driven accommodation of a sufficiently small immediate situation containing just one cat. The bare NP ‘cat’ is not an option in this context because it is not necessarily clear which cat you want. Again, this is compatible with an analysis of bare SG NPs in Guébie as indefinites.

4.3 Bare NPs are indefinite NPs

Turning to the semantics of bare SG NPs in Guébie, a natural solution given their default occurrence in larger and immediate situation uniqueness contexts, cf. (4), (10-b) and (6-a) would be to treat them as unique definites with a covert iota-operator, on a par with bare NPs in Mandarin and Thai (Jenks, 2015; Jenks, 2018), and in line with the analyses of definite bare nouns proposed in earlier work, such as Dayal (2004). A problem with this approach, however, is that the distribution of bare NPs in Guébie differs from that of bare NPs in Mandarin and Thai, cf. Tables 1 and 2 above. In particular, bare SG NPs in Guébie are infelicitous with part-whole bridging or situation-based covariation, unlike in Mandarin/Thai with covert iota-NPs. Conversely, bare SG NPs are licit with producer-product bridging in Guébie, unlike in Mandarin or Thai, where the indexed demonstrative is sometimes required.

Given these distributional differences, we conclude that bare SG NPs in Guébie are not unique definite NPs. Instead, and following up on an analysis in Philipp (2022) for bare NPs in Akan, we will analyse Guébie bare NPs as indefinite NPs that make a rather weak semantic claim on the existence of an individual with the NP-property in question. Crucially, this existence claim is also satisfied with uniquely referring singleton-NPs such as ‘sun’ or ‘president’. In the absence of a plain uniqueness DEF-marker in Guébie, bare indefinite NPs are not blocked from referring to unique referents by an anti-presupposition of non-uniqueness, unlike their indefinite counterparts in English or German, cf. Heim (1991). There is, however, another anti-presupposition triggered by the existence of the overt indexed unique DEF-marker $=a$. This anti-presupposition requires that the referent of the bare NP be non-indexed, i.e. unfamiliar or novel in the context. This directly accounts for the infelicity of Guébie bare SG NPs in anaphoric contexts, perhaps with the exception of the open product-producer puzzle. There hence remain two subcases of the felicitous use of a bare SG NP in Guébie, shown in (24-ab).

- (24) a. non-indexed + singleton NP-restriction (‘sun’, ‘director’, ...): uniqueness DEF
 b. non-indexed + non-singleton NP-restriction (‘dog’, ‘table’): non-specific INDEF

Firstly, bare NPs can be used with novel instances of singleton-referring NPs, as in (4) and (6-a). In such cases, they are surface-equivalent to bona fide uniqueness DEFs even though they come with a different underlying semantics. Secondly, they can be used with novel instances of NPs with non-singleton reference, in which case they receive the default interpretation of a non-specific indefinite NP, also familiar from the discussion of bare NPs in Akan in Philipp (2022). The following example illustrates the use of bare NPs as non-specific indefinites:

- (25) e² ka² no³ gba¹ e² ka³ na-ɓriki^{24.4.1} ɓita^{2.3} sra² ne²
 2SG.NOM IRR do that 2SG.NOM IRR 2SG.POSS-brick.FR house build REL
 ‘If you want to build **a house** with bricks.
 Source: House-building text, spoken by Badiba Olivier Agodio on 8/11/2016

If our analysis is on the right track, it suggests that Guébie lacks access to an ι type-shift altogether.

5 Conclusion

With the lexical entry for the Guébie DEF-marker =*a* in (20), we offer support for the idea that there is a family of indexed definite expressions (Jenks and Konate, 2022), which =*a* is a novel instance of. This conclusion necessitates a revision of the commonly accepted tripartition in (26), based on Ahn (2017), where anaphoric DEFs and demonstratives form the subclass of indexed definites, to the new typology in Table 3. Here, the novel type of DEF-marker is situated with traditional anaphoric DEFs and demonstratives.

- (26) uniqueness DEFs – anaphoric DEFs – Demonstratives

Unique Definites	Indexed Definites		
ι	Anaphoric DEF	Indexed Unique DEF	Demonstratives

Table 3: Definite Expression

Of course, this revision in the typology of DEF-marking has repercussions for cross-linguistic semantic research in definiteness phenomena. Specifically, it necessitates a widening of the empirical search space and the corresponding diagnostic tools.

The difference in =*a*-marking with donkey anaphora in (12), as opposed to the absence of such =*a*-marking with situation-based covariation in (9), furthermore has implications for the formal analysis of donkey anaphora. In particular, Elbourne (2005) argues that donkey anaphora should be analysed on a par with situation-based covariation, namely by a process of situation-binding. In contrast, Schwarz (2009) argues for a different treatment of situation-based covariation, on the one hand, and donkey anaphora, on the other. The different marking patterns observed with situation-based covariation and donkey anaphora in Guébie seem to provide an argument for Schwarz’s

non-unified account, at least as far as Guébie is concerned: The Guébie data suggest that donkey anaphora should receive a dynamic semantic analysis in terms of a DRT-style binding of $=a$'s individual index by some discourse-accessible individual discourse referent (Kamp, 1981, Heim, 1982); cf. Schlenker (2011) and Jenks (2018) for similar conclusions from ASL and Mandarin. The unavailability of bare SG NPs in situation-based covariation moreover provides additional support for our analysis of Guébie bare SG NPs as semantically indefinite. Being indefinite, they do not seem to contribute a (covert) definite determiner with a situation variable to be bound. This would however be required for ensuring uniqueness of the NP-denotation relative to the individual situations quantified over.⁷

Next, the curious absence of demonstratives in Guébie also raises some important questions. In view of the analysis of demonstratives as dimension shifters in Ebert et al. (2020), there is for instance the question of whether Guébie has other indexical elements that could accompany pointing or other manual gestures, such as, e.g., ‘such’, ‘this way’, ‘so’ etc. At the same time, it is well-known from diachronic work that DEF-articles develop from deictic demonstratives via the intermediate step of non-deictic demonstratives, cf. Lyons’s (1999, p. 331) *definiteness cycle*. Relatedly, Simonenko and Carrier (2022) show that the pointing-function of demonstratives is one of the first functions to disappear on the developmental path from demonstratives to definites (cf. also Ahn, 2022). If so, $=a$ may well be in the first stage of its diachronic path from a full-blown demonstrative to a definite determiner, namely at the stage of a *neutral demonstrative* in Lyons’ terminology. A final question is why there is no other functional element in Guébie to take over the lost deictic pointing function of demonstratives, on a par with Akan *saa* in (18), for instance.

Finally, what are we to make of the apparent violation of Diessel’s (1999, p. 1) absolute universal that all languages have demonstratives? In view of the fact that demonstratives constitute a subtype of indexed definite, we propose to maintain a slightly weaker but more precise version of Diessel’s original claim, cf. (27):

(27) All languages have indexed definite NPs.

This version of Diessel’s universal suggests that all languages share the ability to use logical variables in definite expressions to guarantee reference to contextually salient individuals.

References

- Agodio, Olivier et al. (2014). *Guébie Fieldwork Collection, 2014-15*. California Language Archive, Survey of California and Other Indian Languages, University of California, Berkeley, url-<http://dx.doi.org/doi:10.7297/X208639V>.
- Ahn, Dorothy (2017). “Definite and demonstrative descriptions: a micro-typology”. In: *Proceedings of GLOW in Asia xi*.
- (2022). “Indirectly direct: An account of demonstratives and pointing”. In: *Linguistics and Philosophy* 45, pp. 1345–1393.

⁷The use of a bare PL NP with situation-based covariation in (9) remains puzzling to us at present. One possibility would be that such cases in Guébie do not involve situation-binding at all, and that the pairing of churches and priests is achieved by a process of PL-based cumulativity.

- Akumbu, Pius and Peter Jenks (2023). *Demonstratives and definiteness in Babanki*. Handout from presentation at Workshop on Definiteness in the Niger-Congo Noun Phrase, May 2023.
- Arkoh, Ruby and Lisa Matthewson (2013). “A Familiar Definite Article in Akan”. In: *Lingua* 123, pp. 1–30.
- Barlew, Jefferson (2014). “Salience, uniqueness, and the definite determiner -tè in Bulu”. In: *Proceedings of Semantics and Linguistic Theory (SALT)*. Vol. 24, pp. 619–639.
- Bombi, Carla (2018). “Definiteness in Akan (Kwa): familiarity and uniqueness revisited”. In: *Proceedings of Semantics and Linguistic Theory (SALT)*. Vol. 28, pp. 141–160.
- Dayal, Veneeta (2004). “Number marking and (in)definiteness in kind terms”. In: *Linguistics and Philosophy* 27, pp. 393–450.
- Dayal, Veneeta and Li Julie Jiang (2022). “The Puzzle of Anaphoric Bare Nouns in Mandarin: A Counterpoint to Index!” In: *Linguistic Inquiry* 54.1, pp. 147–167.
- Diessel, Holger (1999). *Demonstratives: Form, function and grammaticalization*. Amsterdam: John Benjamins Publishing.
- Ebert, Christian et al. (2020). “Demonstratives as dimension shifters”. In: *Proceedings of Sinn und Bedeutung*. Ed. by Michael Franke et al. Vol. 24, pp. 161–178.
- Elbourne, Paul (2005). *Situations and Individuals*. MIT Press.
- Hawkins, John A. (1978). *Definiteness and Indefiniteness: A Study in Reference and Grammaticality Prediction*. London: Croon Helm.
- Heim, Irene (1982). “The Semantics of Definite and Indefinite Noun Phrases”. PhD thesis. University of Massachusetts, Amherst.
- (1991). “Artikel und Definitheit (Articles and definiteness)”. In: *Semantics: An International Handbook of Contemporary Research*. Ed. by Arnim von Stechow and Dieter Wunderlich. Berlin: deGruyter.
- Jenks, Peter (2015). “Two kinds of definites in numeral classifier languages”. In: *Proceedings of SALT* 25.
- (2018). “Articulated definiteness without articles”. In: *Linguistic Inquiry* 49.3, pp. 501–536.
- Jenks, Peter and Rassidatou Konate (2022). “Indexed definiteness”. In: *Glossa: a journal of general linguistics* 7.1. DOI: <https://doi.org/10.16995/glossa.5717>.
- Jiang, Li Julie (2012). “Nominal phrases and language variation”. PhD thesis. Harvard University.
- (2020). *Nominal arguments and language variation*. Oxford University Press.
- Kamp, Hans (1981). “A theory of truth and semantic representation”. In: *Formal Methods in the Study of Language*. Ed. by J. A. G. Groenendijk et al. Amsterdam: Mathematical Centre, pp. 1–41.
- Löbner, Sebastian (1985). “Definites”. In: *Journal of Semantics* 4, pp. 279–326.
- Lyons, Christopher (1999). *Definiteness*. Cambridge University Press.
- Newman, Paul (2000). *The Hausa Language. An Encyclopedic Reference Grammar*. New Haven: Yale University Press.
- Owusu, Augustina P. (2022). “Cross-categorial definiteness/familiarity”. PhD thesis. Rutgers.
- Philipp, Mareike (2022). “Quantifier Scope Ambiguities in English, German, and Asante Twi (Akan): Structural and Pragmatic Factors”. PhD thesis. Universität Potsdam.
- Sande, Hannah (2020). “Guébie Language Snapshot”. In: *Language Documentation and Description* 19, pp. 35–44.

- Schlenker, Philippe (2011). “Donkey anaphora: the view from sign language (ASL and LSF)”. In: *Linguistics and Philosophy* 34, pp. 341–395.
- Schwarz, Florian (2009). “Two Types of Definites in Natural Language”. PhD thesis. UMASS.
- Simonenko, Alexandra and Anne Carlier (2022). *Demonstratives and strong definites: two ways to point to a situation*. Handout from presentation at 2nd DFG-Network Meeting on Definiteness across Domains, December 2022.
- Zimmermann, Malte (2008). “Quantification in Hausa”. In: *Quantification: Universals and Variation*. Ed. by Lisa Matthewson. Bingley: Emerald.

From a cross-linguistic perspective, intervention effects have been observed in a variety of languages among which are Korean, Hindi, Turkish (Beck 1996), English, French, Japanese (Pesetsky 2000), Russian, Palestinian Arabic, Samoan, Yoruba (Howell et al. 2022), and many more. This paper adds one more language, namely Tundra Nenets, to the cross-linguistic landscape.

The paper is structured as follows: Section 2 addresses the theoretical background, in particular the distinguished variables framework and the cross-linguistic picture. Section 3 starts with a short introduction to the Nenets language before delving into the prerequisites for testing intervention effects in TN. These prerequisites include focus marking, question formation, sensitivity to alternatives, and association of the operators Q and squiggle, \sim , with their alternative-introducing items at a distance. The paper then proceeds to test the relevant intervention effects in TN in section 4. Finally, section 5 contains a discussion of the findings and their broader implications.

2 Theoretical Background

2.1 Distinguished Variables Framework

In the *distinguished variables* framework (cf. Kratzer 1991, Wold 1996, Beck 2006, 2016), alternatives are introduced into the semantics by focus and question words in form of a distinguished variable, i.e. the *wh*-pronoun *who* in (4-a) and the focused subject *Olga* in (4-b).

- (4) a. **Who** does fieldwork?
 b. Only **Olga_F** does fieldwork.

Both introduce a distinguished variable which is then interpreted by the distinct variable assignment function h . Distinguished variables introduced by focus are bound by a single focus-evaluating operator, the squiggle-operator \sim , going back to Rooth (1992) in (5), while those introduced by a *wh*-pronoun are bound by the Q -operator in (6) (from Howell et al. 2022 who adapted the lexical entry from Beck 2016). Association with focus is mediated by \sim , which introduces a presupposition on the value of a free contextual variable C . This amounts to Rooth's original \sim -operator spelt out in a framework that uses distinguished variables. Like in Rooth's original proposal, the \sim restricts the value of a free variable C to a subset of this set of alternatives and results in an expression where all distinguished variables occurring in its scope are unavailable for binding by higher operators.

(5) **Squiggle operator, \sim (unselective):**

If $\alpha = [\sim_i C\beta]$, then for any g, h :

$\llbracket \alpha \rrbracket^g$ is only defined if $g(C) \subseteq \{ \llbracket \beta \rrbracket^{g, h[x/i]} \mid h \text{ a total distinguished variable assignment function} \}$.

Then, $\llbracket \alpha \rrbracket^g = \llbracket \beta \rrbracket^g$ and $\llbracket \alpha \rrbracket^{g, h} = \llbracket \beta \rrbracket^{g, \emptyset}$.

(6) **Question operator, Q (selective)**

If $\alpha = [Q_i\beta]$, then for any g, h and the semantic type τ determined by i :

$\llbracket \alpha \rrbracket^g = \{ \llbracket \beta \rrbracket^{g, \emptyset[x/i]} \mid x \in D_\tau \}$

$\llbracket \alpha \rrbracket^{g, h} = \{ \llbracket \beta \rrbracket^{g, h[x/i]} \mid x \in D_\tau \}$

In the case of questions, the set resulting from binding of distinguished variables by the Q -operator becomes the question meaning. The LF for the sentence (4-a) (repeated in (7-a)) is in (8), the truth conditions are in (7-c).

- (7) a. **Who** does fieldwork?
 b. $\llbracket who_{i,e} \rrbracket^g$ undefined;
 $\llbracket who_{i,e} \rrbracket^{g,h} = h(i)$
 c. $\llbracket (7-a) \rrbracket^{g,h} =$
 $\{ \llbracket [\lambda w. [who_i [w \text{ do. fieldwork}]]] \rrbracket^{g,h[x/i]} | x \in D_e \}$
 $= \{ \lambda w. h[x/i](i) \text{ do. fieldwork in } w | x \in D_e \}$
 $= \{ \lambda w. x \text{ do. fieldwork in } w | x \in D_e \}$
 $= \{ \lambda w. \text{ Julia does fieldwork in } w,$
 $\lambda w. \text{ Bayan does fieldwork in } w, \dots \}$
- (8)
-

Association of focus-sensitive particles like *only* with focus is mediated by \sim (indirect association with focus — going back to Rooth 1992): The free-variable sister to \sim , whose value is constrained by the presupposition introduced by \sim , is co-indexed with the first argument of *only*. This means that focus-sensitive particles always come with a \sim . The LF for (4-b) (repeated in (9-a)) is in (9-b), the definedness condition is in (9-c). A simplified lexical entry for the exclusive is in (9-d) and finally, the truth conditions of the sentence under both g and h functions are in (9-e).

- (9) a. **Only Olga_F** does fieldwork.
 b. $\llbracket t [\langle st, t \rangle \text{ only}_{w@} C_{1, \langle st, t \rangle}] [[\sim C_{1, \langle st, t \rangle}] [\langle s, t \rangle \text{ Olga}_{i,e} [\text{do fieldwork}]_{\langle e, \langle s, t \rangle \rangle}]] \rrbracket$
 c. **presupposition:** $\llbracket g(1, \langle st, t \rangle) \rrbracket = \{ p : \exists x \in D_e [p = \lambda w. \text{do. fieldwork}_w(x)] \}$
e.g. $\{ \lambda w. \text{do. fieldwork}_w(\text{Petya}), \lambda w. \text{do. fieldwork}_w(\text{Katya}), \lambda w. \dots \}$
 d. $\llbracket \text{only}_{(\text{simplified})} \rrbracket = \lambda w. \lambda C_{\langle st, t \rangle}. \lambda p_{\langle s, t \rangle}. \forall q [C(q) \ \& \ q \neq p \rightarrow \neg q(w)]$
 e. $\llbracket \dots \rrbracket^g = \lambda w. \text{do. fieldwork}_w(\text{Olga})$
 $\llbracket \dots \rrbracket^h = \lambda w. \text{do. fieldwork}_w(h(i, \langle e \rangle))$

Crucially, the binding properties of Q and \sim differ in that Q **binds a distinguished variable selectively**, while \sim unselectively binds everything in its scope.

2.2 Cross-linguistic Picture

Howell et al. (2022) investigated cross-linguistic variation in the grammar of alternatives with a specific focus on intervention effects. They conducted an in-depth cross-linguistic study of five languages. Assuming that both the selective and unselective versions of Q and \sim are available to the grammar, the possible patterns in Table 1 should hypothetically exist in the world's languages.

pattern # 1 unselective \sim selective Q	pattern # 3 unselective \sim unselective Q
pattern # 2 selective \sim selective Q	pattern # 4 selective \sim unselective Q

Table 1: Hypothetical Typology of Alternative Evaluation

Somewhat surprisingly, Howell et al. (2022)'s cross-linguistic survey did not find the theoretically possible variation illustrated in Table 1 with respect to the selectivity properties of \sim and Q . In all the languages investigated, focus-sensitive items gave rise to intervention effects, pointing towards an unselective \sim operator. On the other hand, association with focus and Q was possible across the Q -operator in the majority of languages tested. The results are summarized in Table 2. The authors propose two semantic universals in (10) which capture the cross-linguistic pattern they found.

	\sim	Q	pattern #
English	unselective	selective	1
German	unselective	selective	1
Palestinian Arabic	unselective	selective	1
Russian	unselective	selective	1
Samoan	unselective (?)	selective	1
Turkish	unselective	selective	1
Yoruba	unselective	selective	1

Table 2: Summary of Results (Howell et al. 2022)

- (10) a. **Universal 1: Unselective Squiggle**
Universal 1a: Association via Squiggle
Focus evaluation is always mediated by the same focus-evaluating operator \sim .
Universal 1b: Unselective Association
In all languages, \sim is an unselective alternative-evaluating operator.
- b. **Universal 2: Selective Q**
In all languages, the Q -operator associates with *wh*-items or disjunction in its scope selectively.

Looking at the results, the question arises whether this lack of variation is just an artifact of the languages that were under investigation, i.e. a mere coincidence, or whether there is more to it, like a cross-linguistic universal as suggested by Howell et al. (2022). This leads me to the following research question in (11):

- (11) Do we find the predicted lack of variation with respect to the selectivity properties of \sim and Q in Tundra Nenets?

3 Prerequisites for Testing Intervention Effects in TN

This section begins by introducing some general facts about the TN language in general and its grammar in particular. It then proceeds to establish the prerequisites for testing the intervention configurations.

Nenets, or Yurak, is a Samoyedic language from the Uralic language family. Nenets has two dialects: the Tundra Nenets and the Forest Nenets dialect. According to Ethnologue (Eberhard, Simons and Fenning (eds.) 2023), there are over 20,000 Tundra Nenets speakers and about 1500 Forest Nenets speakers with the total of 49,600 (2020 census) Nenets people (ethnic population). The language status is categorized as 6b, i.e. threatened. I worked with consultants in Arkhangelsk and Naryan-Mar in the Nenets Autonomous Okrug (NAO). I also had the chance to work with Nenets students of the Herzen University in Saint Petersburg. The geography of the Nenets language along with the loci of my fieldwork is shown in the following map.



The shading indicates the regions where Nenets is spoken.

Map: Kinga Kuti, Polina Berezovskaya • Source: Ethnologue • Created with Datawrapper

Figure 1: Geography of Nenets and Loci of my Fieldwork

In total, I worked with 19 language consultants who spoke different subdialects of TN: the Kanine, the Kolguyev, the Yamal, as well as the those of Malaya and Bol'shaya Zemlya. These subdialects are mutually intelligible even if they show some variation, especially in the phonology, but also in the lexicon.

Let us now turn to the most relevant grammatical features: TN is a highly agglutinative language. The two major word classes are verbs and nouns. Adjectives, adverbs and postpositions also exist, but often exhibit certain properties of nouns. In terms of syntax, the basic word order is SOV with verb finality being a rigid constraint as shown in ex. (12). Variation is present in the placement of the subject vs. the object. TN is quite consistently head-final: It is a verb-last language, it has postpositions, etc.

- (12) **word order in Nenets in a regular transitive sentence:**
 (Time adverbial)-subject NP-(place adverbial)-indirect object NP
 -object NP-(manner adverbial)-verb.
 (cf. e.g. Salminen 1998, Nikolaeva 2014: 214)

In order to answer the research question in (11), I tested intervention effects by \sim and Q in the following configurations in (13) from Howell et al. (2022):

- (13) INTERVENTION CONFIGURATIONS
- a. $[\sim [\dots [Q_i [\dots wh_i \dots F]] \dots]]$
 Example: *I only asked who likes durian_F.*
 (Association with focus across Q)
- b. $[Q_i [\dots [Q_{ii} [\dots wh_{ii} \dots wh_i \dots]]] \dots]]$
 Example: *Who knows where we bought what?*
 ‘For which person-thing pairs $\langle x, y \rangle$: x knows where we bought y?’
 (Association with Q across Q . (Baker 1968/70))
- c. $*[Q_i [\dots [\sim [\dots wh_i \dots F \dots]]] \dots]]$
 Example: **Which student did only Karl_F recommend which book to?*
 (Association with Q across focus = focus intervention)
- d. $*[\sim [\dots [\sim [\dots F \dots F \dots]]] \dots]]$
 Example: *?I also only introduced Marilyn_F to Ted_F.*
 (Association with focus across focus = focus intervention)

Before I was able to test these particular configurations in TN, the following prerequisites had to be checked first:

- i focus marking
- ii question formation
- iii sensitivity to alternatives, and
- iv (im)possibility of long-distance association of Q and \sim .

The importance of first clarifying all of these prerequisites should become clear with the following example: In *wh*-fronting languages like Russian, a configuration where the Q -binder is separated from its *wh*-pronoun is hard to come by. Berezovskaya and Howell (2020) show that such a configuration is possible in Russian, but only in embedded questions. This kind of insight could

not be arrived at without first checking how question formation works, whether (long-)distance association of *Q* with the *wh*-item is possible, etc. Therefore, it is important to first establish whether TN is a *wh*-in-situ language or not, how focus is marked, which are the focus-sensitive items and whether *Q* and \sim can associate with the alternative-introducing item at a (long-)distance.

3.1 Focus Marking in Nenets = Prerequisite (i)

Introduction of alternatives in Nenets happens via phonological focus marking, but also by *wh*-items. There is no dedicated focus marker. However, as I found in my fieldwork, optional focus marking is possible with the particle *-ri*, which Nikolaeva (2014) calls the ‘limitative’ with “the most frequent meaning” being ‘only’ (Nikolaeva 2014: 124). According to Nikolaeva, this suffix occurs on nouns, pronouns, verbs, adverbs and postpositions. For nouns and adjectives, it precedes inflectional morphology, i.e. case, number and possessive affixes.² I illustrate the occurrence of *-ri* on a noun with an exclusive meaning in (14). In a scenario where the only individual working is Masha, (14) is appropriate with *-ri* on the focused constituent (Masha in this case)³.

- (14) Masha-ri manzara.
 Masha-RI work>3.SG
 ‘Only Masha works.’

Based on most of Nikolaeva’s data and data like (14), at first glance one could take *-ri* to be a phrasal exclusive particle. However, in (15) and (16) (with picture contexts taken from Renans, Zimmermann, and Greif 2011), person B can contradict person A with the additive particle *ɲobtaremh* or *tamna*.

- (15) a. A: Valakada Masha-ri
 A: EXCL Masha-RI
 banan-m ɲavorɲa.
 banana-ACC eat>3.SG
 A: ‘Only Masha eats a banana.’
 b. B: Ni, Anja-ri **ɲobtaremh**
 B: no, Anya-RI ADD
 banan-m ɲavorɲa.
 banana-ACC eat>3.SG
 B: ‘No, Anya also eats a banana.’



²Interestingly, free choice pronouns can be derived by adding *-ri* to question words, e.g. *xib'a-ri* = ‘whoever’.

³Abbreviations in glosses: ACC-accusative, ADD-additive particle, DALAT-dalative, DUR-durative, EXCL-exclusive particle, FEM-feminine, IMPF.PART-imperfective particle (Nikolaeva 2014), INSTR-instrumental, INTER-interrogative suffix, LOC-locative, PL-plural, POSS-possessive, SG-singular.

A note on romanization: In Nenets, the Cyrillic script with some extra signs is used. For the transliteration in my glosses, I am using the orthographical representation by Décsy (1966), pp. 9-11.

- (16) a. A: Zhanna valakada ŋamded-m
 A: Zhanna EXCL flower-ACC
 n'ah-mbi.
 hold-DUR>3.SG
 A: 'Zhanna is only holding a
 flower.'
- b. B: Niŋov! Pyda **tamna**
 B: no! She ADD
 xniska-ri-m n'ah-mbi.
 book-RI-ACC hold-DUR>3.SG
 B: 'She also holds a book.'



An exclusive reading for the sentences is implausible. This behavior of *-ri* can, however, be explained if we assume that the additive particle associates with the focused subject marked by *-ri*. This makes the exclusive reading implausible. This also illustrates that the suffix is not itself an exclusive focus-sensitive particle. I conclude instead that *-ri* is an optional focus marker in TN.

3.2 Question Formation in Nenets = Prerequisite (ii)

In TN questions, the basic word order SOV is mostly preserved. An example for a subject question is in (17-a) and an object question in (17-b). In both examples, the canonical SOV verb order is preserved. TN also has multiple questions in which the word order is preserved as well, cf. the examples in (18).

- (17) a. Xib'a lavka-n xaja?
 who shop-DALAT go>3.SG
 'Who goes to the shop?' word order: SV
- b. Masha xib'a-m xa-sa?
 Masha who-ACC call-INTER>3.SG
 'Whom did Masha invite?' word order: SOV
- (18) a. Xib'a ŋamge-m vade-sa?
 who what-ACC say-INTER>3.SG
 'Who said what?' word order: SOV
- b. Xib'a xan'ana il'e?
 who where live>3.SG
 'Who lives where?' word order: S place adverbial V

This and other similar data from my fieldwork lead me to conclude that TN is a *wh*-in-situ language (but cf. Nikolaeva 2014 and Mus 2022 for discussion). In addition, the existence of multiple questions provides the ground for testing sentences with more than one alternative-introducing item in one sentence. This is relevant for the intervention configurations that will be addressed in section 4.

3.3 Sensitivity to Alternatives in Nenets: Exclusives = Prerequisite (iii)

Since we established that the suffix *-ri* is an optional focus marker and not a focus-sensitive particle, we are so far left without any focus-sensitive particle in TN. A promising candidate is the exclusive *valakada*. It is translated as “only” (Russian *tol’ko*) in the dictionary by Tereschenko (1989). There are some speakers who don’t use *valakada* as an exclusive particle, but rather as an exclamative expressing “Enough!”. This is dependent on the speaker’s subdialect. For those speakers who use *valakada* as an exclusive particle, it can associate with subject focus as in (19), with object focus as in (20-a) and with verb focus as in (21-a).⁴

- (19) Vanja-vah valakada xniska-m-da temda.
 Vanya-POSS.1.PL EXCL book-ACC-POSS.3.SG buy>3.SG
 ‘Only our Vanya_F bought his book.’ (Olya and Senya didn’t buy any book).
- (20) Context: Vanya spent a lot of time choosing for a long time between *War and Peace*, *Anna Karenina* and *Resurrection* by Lev Nikolaevich Tolstoy. In the end...
- a. Pyda valakada Anna Karenina-m tolz’o xniska-m temda.
 he EXCL Anna Karenina-ACC to.read book-ACC buy>3.SG
 ‘He only bought Anna Karenina_F.’
- (21) Context: The book lay at Vanya’s home for quite a while. He hasn’t read it, not even browsed through it. In the end...
- a. Pyda valakada Anna Karenina-m tolz’o xniska-m temda.
 he EXCL Anna Karenina-ACC to.read book-ACC buy>3.SG
 ‘He only bought_F Anna Karenina. (He didn’t read it, or do something else with it.)’

Analyzing the suffix *-ri* as an optional focus marker makes the prediction that we should be able to find cases where *valakada* and *-ri* co-occur. This is in fact borne out, as the following example in (22)⁵ illustrates.

- (22) a. Pyda **valakada** Anna Karenina-**ri**-m tolz’o xnisk-**ri**-m temda.
 he EXCL Anna Karenina-RI-ACC to.read book-RI-ACC buy>3.SG
 ‘He only bought Anna Karenina_F’.
- b. Vanja-**ri**-vah **valakada** xniska-m-da temda.
 Vanya-RI-POSS.1.PL EXCL book-ACC-POSS.3.SG buy>3.SG
 ‘Only our Vanya_F bought the book.’

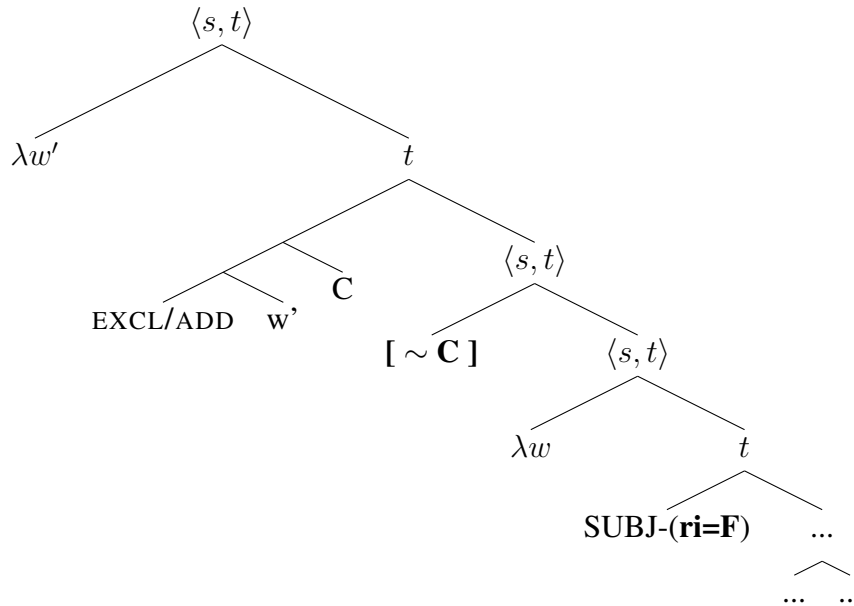
The strategy of using both *valakada* and *-ri* on the focused constituent was even the preferred strategy for some of my consultants. Note how this, again, supports my proposal that the suffix *-ri* is an optional focus marker. It indicates the presence of the squiggle operator higher up in the structure. I assume the following Logical Form in (23) for Nenets focus marking. Focus is evaluated via squiggle higher up in the tree, with *-ri* optionally marking the focus on the focused

⁴There also seems to be variation in that the association at a distance is only available for a subset of those speakers who use *valakada* as an exclusive focus-sensitive particle.

⁵In (22-a), *-ri* even appears twice, on the incorporated noun ‘book’ and on the title, i.e. *Anna Karenina*. Nikolaeva (2014) calls that “limitative concord.”

constituent, in this case the subject.

(23)



3.4 (Long-)Distance Association = Prerequisite (iv)

The final box to tick off with respect to the prerequisites for testing intervention in TN is whether association is possible at a distance between the operators Q and \sim with their respective alternative-introducing items, i.e. wh -items for Q and focused constituents for \sim . The example in (24), as well as previous examples in (17-b) and (18) all illustrate that the Q -operator can be detached from its wh -words, i.e. they can associate at a distance with the Q -operator.

- (24) Xib'a xib'andun'a xada-na t'eb'er-sa?
 Who with.whom nail-LOC strike-INTER>3.SG
 'Who fought with whom?'

[$Q_{i,ii} \dots wh_i \dots wh_{ii}$]

What about the association with the squiggle operator across a distance? This can be tested with the exclusive *valakada* that was introduced in section 3.3. If we look at its selectional restrictions, *valakada* likes to be adjacent to the constituent it associates with. Association at a distance is however possible, like the following examples show.

- (25) Context: The cook has decided to poison his guests (because he owes them big sums of money and is afraid of revenge). He decided to put poison into the soup. He didn't realize that the poison also got into the meat and the potatoes, that constitute the main dish.
- a. η avar piri-ba-da **valakada** ibidor η a pyda ja-m vomda.
 food cook-DUR-IMPF.PART EXCL think>3.SG he soup-ACC spoil>3.SG
 'The cook only thinks he poisoned the soup.' (But in fact, he poisoned the meat and the potatoes, too.)

- b. η avar piri-ba-da ibidor η a pyda ja-**ri**-m vomda.
 food cook-DUR-IMPF.PART think>3.SG he soup-**RI**-ACC spoil>3.SG
 ‘The cook (lit. the food cooking person) thinks he poisoned only the soup.’

In (25-a), the exclusive particle *valakada* associates with the focused element *jam* (‘soup’) long-distance. Another example for this possibility is in (21-a). For those speakers who don’t use *valakada* as an exclusive marker, the strategy in (25-b) is the preferred one. For all examples, this shows the configuration: [\sim ... **F_i**]. I conclude that association of both the *Q*-operator and the squiggle with the alternative-evaluating item at a distance is possible in TN. I also tested other alternative-evaluating operators such as the additive particles *η obtaremh* and *tamna* which can be seen in examples (15) and (16). However, due to too little and partly inconclusive data, I am focusing on the exclusive particle *valakada*. The following Table 3 summarizes the prerequisites section.

Questions		
<u>simple wh-question</u>	<u>multiple wh-question</u>	<u>alternative questions⁶</u>
yes	yes	yes
(Focus-sensitive) particles		
<u>particle</u>	<u>focus sensitive?</u>	<u>Distance association at LF?</u>
<i>valakada</i> (EXCL)	yes	yes
<i>ηobtaremh</i> (ADD)	yes	yes
<i>tamna</i> (ADD)	?	?

Table 3: Summary of the Prerequisites in TN

4 Testing Intervention Effects in TN

We are finally equipped with all the tools for testing the intervention configurations in TN. The working hypothesis for TN is in (26):

- (26) **Working Hypothesis:** Evaluation of alternatives in TN happens via a selective \sim , and a selective *Q*, i.e. TN is a pattern #2 language from Table 1.

If the hypothesis is confirmed and Nenets is indeed a pattern #2 type language, then the potential universal **1b** in (10-a) regarding squiggle from Howell et al. (2022) is challenged. I will go through all the configurations and discuss each of them in turn.

⁶I am including alternative questions in the table for the sake of completeness. However, it is not clear to me what the exact correlate of the natural language disjunction is in TN in the first place. The most likely candidate is *nibtja*, as in the example below. I also have too little data to test intervention in these questions.

- (i) Pydar nibtja halja-m, η amza-n xarva-n?
 You or fish-ACC meat-DALAT want-2.SG
 ‘Do you want fish or meat?’

For discussion of how intervention effects can be tested in alternative questions, see p. 9 ff. in Howell et al. (2022).

Selectiveness of Q /non-intervention by Q . The first configuration I will address is in (27) repeated from above (in (13-b)):

(27) [Q_i [... [Q_{ii} [... wh_{ii} ... wh_i ...]] ...] (Association with Q across Q)

I tested Baker (1968) ambiguities for sentences like ‘Who knows where we bought what?’ in order to check for the selectiveness of Q . The second reading below (Reply B) is only available with a selective Q .

(28) Xib’a teneva xan’ana man’ah ŋamge-m temdasa-vah?
 who know>3.SG where we what-ACC buy-1.PL
 ‘Who knows where we bought what?’

Reply A:

a. Vanja teneva xan’ana man’ah pany-m, nara-m, temda-va-c’h.
 Vanya know>3.SG where we fur.coat-ACC shelf-ACC buy-1.PL-PAST
 ‘Vanya knows where we bought the fur coat, the shelf, etc.’

Reply B:

b. Vanja teneva xan’ana man’ah pany-m temda-va-c Petja
 Vanya know>3.SG where we fur.coat-ACC buy-1.PL-PAST, Petya
 teneva xan’ana man’ah nara-m temda-va-c’,...
 know>3.SG where we shelf-ACC buy-1.PL-PAST...
 ‘Vanya knows where we bought the fur coat, Petya knows where we bought the shelf, etc.’

The Baker diagnostics worked robustly for several Nenets speakers. This already points to a selective Q . The next example instantiates the configuration from (13-a) repeated in (29). In (30-a), *valakada* is above the embedded question indicating the presence of \sim higher up in the structure.

(29) [\sim [... [Q_i [... wh_i ... F]] ...] (Association with focus across Q)

(30) Context: Irina is always very well informed about everything that is going on in the village. She has information on who is dating whom, who is new, etc. Soon after three new girls (Lena, Masha and Vika) arrive at the village, people start asking about whether Irina knows whether any of the new girls is in love with somebody. However, Irina only knows whom Masha loves.

a. Irina valakada teneva xib’a-m Masha-vah men’e.⁷
 Irina EXCL know>3.SG who-ACC Masha-POSS.1.PL love>3.SG
 ‘Irina knows whom only our Masha_F loves.’

This example also speaks in favor of a selective Q .

⁷Interestingly, my consultants often offered this possessive marking on the nominal focused constituent in my examples. This amounts to saying something like: “our Masha...”. This marking indicates familiarity with the introduced person in the context and curiously coincides with focus.

Non-intervention by \sim / Selectiveness of \sim . Remember that examples like (17) and (18) from above show that the base order, i.e. SOV, is also the preferred one in questions. Using the knowledge that Nenets is a *wh*-in-situ language and adding the phrasal *-ri* to the picture, I tested the configuration in (13-c) repeated in (31) both with *-ri* and with *valakada*. The only difference in (31) is that the two alternative-introducing elements (*wh* and F) are turned around, which does not affect the overall configuration since only the scope relations matter.

(31) $*[Q_i [\dots [\sim [\dots F \dots wh_i \dots]] \dots]]$

(Association with *Q* across focus = focus intervention)

(32) Context: There is a person that only Masha invited to the party. I am curious to know who that person is and ask:

- a. Masha-ri xib'a-m xansa?
 Masha-RI who-ACC called
 'Whom did only Masha invite?'

Since I established that *-ri* is just an optional focus marker and not a focus-sensitive particle, this configuration is not a suitable intervention configuration, since we cannot be sure about the position of the squiggle. However, this example can be replicated with the exclusive *valakada*.

(33) Valakada Masha-ri xib'a-m xansa?
 EXCL Masha-RI who-ACC called
 'Who did only Masha invite?'

The example in (33) is also fine when we leave out the focus marking *-ri*. However, leaving *-ri* on the focused subject was the preferred variant. We need to assume that instead of intervening between *Q* and the *wh*-pronoun, \sim binds only the distinguished variable introduced by focus. 'Masha' is bound by \sim , however *Q* wants to bind its *wh*-pronoun, too. The solution in this case is that \sim binds only the focused constituent 'Masha' selectively, while the *Q*-operator selectively binds the *wh*-pronoun *xib'am*. Since this configuration is good with the exclusive particle *valakada*, we have to assume a selective squiggle here! Otherwise, the question should be infelicitous.

Additional data in favor of non-intervention by \sim comes from multiple focus configurations. While Howell et al. (2022) leave out multiple focus constructions due to the fact that the status of such constructions in the literature is highly controversial (Wold 1996, Rooth 1992 and others report acceptable judgments, while experimental results from Beck and Vasisht (2009) lead us to think of those as degraded in English), I still tested these constructions in TN. An example that instantiates the potential intervention configuration in (13-d) repeated below in (34) is in (35). The sentence in (35-a) is meant as a contradiction to what is established in the context, namely that the only person I introduced Masha to is Petya.

(34) $*[\sim [\dots [\sim [\dots F \dots F \dots]] \dots]]$

(Assoc. with focus across focus = focus intervention)

- (35) Context: I only introduced Masha_F to Petya.
- a. Man' ğobtaremh valakada Masha-ri-m Tolja-ri-n toromdav.
 I ADD EXCL Masha-RI-ACC Tolya-RI-DALAT introduced
 'I also only introduced Masha_F to Tolya_F.'
 reading: 'Another person who I introduced only Masha to is Tolya.'

The higher \sim needs to associate with the lower focus that is c-commanded by the closer (second) \sim , which comes with its own (higher) focus. The sentence in (35-a) on the intended interpretation thus requires the configuration in (36).

- (36) [\sim_1 ... [\sim_2 [X ... F2 ... F1 ...]]]

This configuration requires a selective squiggle for which a meaning rule is provided in (37).

- (37) **Wold's selective \sim :**
 If $\alpha = [\sim_i \text{ c } \beta]$, then for any g,h:
 $\llbracket \alpha \rrbracket^g$ is only defined if $g(C) \subseteq \{\llbracket \beta \rrbracket^{g,h[x/i]} \mid x \in D\}$.
 Then, $\llbracket \alpha \rrbracket^g = \llbracket \beta \rrbracket^g$ and $\llbracket \alpha \rrbracket^{g,h} = \llbracket \beta \rrbracket^{g,h}$.

More data needs to be elicited for Nenets in order to clarify whether the absence of intervention effects persists with other focus-sensitive particles. For now, the following table in 4 summarizes the results from this section. All of the tested configurations yielded grammatical and interpretable sentences, which is indicated by a checkmark in the table.

CONFIGURATIONS			
Intervention by \sim ?		Intervention by Q ?	
Q across \sim	\sim across \sim	Q across Q	\sim across Q
✓	✓	✓	✓

Table 4: Summary of the Intervention Configurations in TN

The hypothesis from (26) is therefore confirmed: According to my data, TN has not only a selective Q like the languages in the Howell et al. (2022) language sample, but also a selective \sim .

5 Discussion

As illustrated in the previous sections, I did not find any intervention effects in the configurations from Howell et al. (2022). Contra the expectation from languages exemplified in Howell et al. (2022), TN seems to be a language with both a selective Q and a selective \sim . I am adding TN to the list of languages from Howell et al. (2022) in Table 5.⁸

⁸The question marks in the squiggle and the pattern column mean that more focus-sensitive items and more data in general should be tested to see whether the pattern I show in this paper persists.

	~	<i>Q</i>	pattern #
English	unselective	selective	1
German	unselective	selective	1
Palestinian Arabic	unselective	selective	1
Russian	unselective	selective	1
Samoan	unselective (?)	selective	1
Turkish	unselective	selective	1
Yoruba	unselective	selective	1
Tundra Nenets	selective (?)	selective	2?

Table 5: Summary of Results by Howell et al. (2022) with TN Included

The question that arises is whether TN generally does not display intervention effects or whether we don't see them just in the configurations investigated in this paper. Intervention by negation or a negative quantifier is a classic case of intervention from the literature. Interestingly, I did find intervention effects with a negative quantifier in TN, as shown in (38).

- (38) a. Xib'a-m xan'ana Petja xosa?
 who-ACC where Petya met
 'Who did Petya meet where?'
 b. *Xib'a-mh xan'ana xib'axart xosa??
 who-ACC where nobody met
 Intended: 'Whom did nobody meet where?'

TN is not the only candidate where intervention is possible with some, but not all interveners. Eilam (2011) observes that in Amharic there is intervention with Negative Polarity Items, but not with focus-sensitive particles. Firstly, this means that different interveners should not be treated on a par, e.g. negative quantifiers or negative polarity items are different creatures than the interveners we are dealing with in this paper. Secondly, this also leads me to a broader theoretical question. The theoretical underpinning of the present, as well as Howell et al. (2022)'s paper is based on Beck's (2006) view. However, we should not forget that there are also other frameworks on the market. Li and Law (2016), for instance, call Beck's view the "reductionist approach", because it brings together Alternative Semantics by Hamblin (1973) and Focus Semantics by Rooth to explain focus intervention effects by treating focus and *wh*-questions on a par. Beck's approach perfectly explains such phenomena as question/answer congruence in English. But what if there are cross-linguistic differences in this domain? What if we do not see any intervention in TN with the focus/*wh*-configurations because at least in TN these are different creatures that do not interact in the same way as they do in English and possibly other languages? This could explain the lack of intervention we are seeing in TN. However, at this point, this idea remains just a musing without a clear spell-out and must be left for future research.

Another issue needs to be addressed here, namely that theoretically, an alternative explanation of the data needs to be considered. The lack of intervention effects that we are witnessing in TN might be a result of covert focus and *wh*-movement (cf. Karttunen 1977, Drubig 1994, Erlewine and Kotek 2017 a.m.o.) rather than the selectivity properties of alternative-evaluating operators.

However, a strong argument against this alternative account is that multiple *wh*-questions in TN do not seem to exhibit superiority effects⁹ (cf. Pesetsky 2000). This suggests that there is no covert *wh*-movement strengthening the claim of this paper.

As a result, TN represents a very likely candidate for a pattern #2 language. Now, this questions the universal 1b (cf. (10) above) from Howell et al. (2022) that \sim unselectively evaluates all alternatives in its scope. The Nenets data points to the existence of systematic variation in the mechanics of alternative evaluation. Differences in the interaction of alternative-evaluating operators uncover this variation in intervention effects. While in language after language, e.g. in English, German, Korean, Russian etc. we find the theoretically predicted selective Q and unselective \sim , Nenets adds a new language to the picture that does not align with the pattern attested so far.

References

- Baker, Carl Leroy (1968). “Indirect Questions in English”. PhD thesis. University of Illinois at Urbana-Champaign.
- Beck, Sigrid (1996). “Wh-Construction and Transparent Logical Form”. PhD thesis. Eberhard Karls Universität Tübingen.
- (2006). “Intervention Effects Follow from Focus Interpretation”. In: *Natural Language Semantics* 14.1, pp. 1–56. DOI: <https://doi.org/10.1007/s11050-005-4532-y>.
- (2016). “Focus Sensitive Operators”. In: *The Oxford Handbook of Information Structure*. Ed. by Caroline Fery and Shinichiro Ishihara. Oxford University Press, pp. 227–250. DOI: <https://doi.org/10.1093/oxfordhb/9780199642670.013.007>.
- Beck, Sigrid and Shravan Vasishth (2009). “Multiple Focus”. In: *Journal of Semantics* 26.2, pp. 159–184. DOI: <https://doi.org/10.1093/jos/ffp001>.
- Berezovskaya, Polina and Anna Howell (2020). “(No) Variation in the Grammar of Alternatives: Intervention Effects in Russian”. In: *Proceedings of Formal Approaches to Slavic Linguistics (FASL) 26*, pp. 1–19.
- Décsy, Gyula (1966). *Yurak Chrestomathy*. Bloomington: Indiana University Publications.
- Drubig, Hans Bernhard (1994). *Island Constraints and the Syntactic Nature of Focus and Association with Focus*. Universitäten Stuttgart und Tübingen in Kooperation mit der IBM Deutschland GmbH.
- Eberhard, David M., Gary F. Simons, and Charles D. Fennig (eds.) (2023). *Ethnologue: Languages of the World*. SIL International. URL: <http://www.ethnologue.com>.
- Eilam, Aviad (2011). “Explorations in the Informational Component.” PhD thesis. University of Pennsylvania.
- Erlewine, Michael Yoshitaka and Hadas Kotek (2017). “Movement and Alternatives don’t Mix: Evidence from Japanese”. In: *Proceedings of the 21st Amsterdam Colloquium*, pp. 245–254.
- Hamblin, C. L. (1973). “Questions in Montague English”. In: *Foundations of Language* 10.1, pp. 41–53. ISSN: 0015900X. URL: <http://www.jstor.org/stable/25000703> (visited on 11/19/2023).

⁹This is an observation both by Nikolett Mus’ (p.c.) from her fieldwork and me in my own fieldwork.

- Howell, Anna, Vera Hohaus, Polina Berezovskaya, Konstantin Sachs, Julia Braun, Şehriban Durmaz, and Sigrid Beck (2022). “(No) Variation in the Grammar of Alternatives”. In: *Linguistic Variation* 22.1, pp. 1–77. DOI: <https://doi.org/10.1075/lv.19010.how>.
- Karttunen, Lauri (1977). “Syntax and Semantics of Questions”. In: *Linguistic and Philosophy* 1, pp. 3–44. URL: <https://doi.org/10.1007/BF00351935>.
- Kratzer, Angelika (1991). “The Representation of Focus”. In: *Semantics: An international handbook of contemporary research*, pp. 825–834.
- Li, Haoze and Jess H.-K. Law (2016). “Alternatives in Different Dimensions: A Case Study of Focus Intervention”. In: *Linguistics and Philosophy* 39, pp. 201–245. DOI: <https://doi.org/10.1007/s10988-016-9189-0>.
- Mus, Nikolett (2022). “Syntax of Multiple Questions in Tundra Nenets”. In: *Studia Uralo-Altaica* 56, pp. 127–138. URL: <https://doi.org/10.14232/sua.2022.56.127-138>.
- Nikolaeva, Irina (2014). *A Grammar of Tundra Nenets*. Berlin, Boston: de Gruyter Mouton, pp. –. DOI: <https://doi.org/10.1515/9783110320640>.
- Pesetsky, David (2000). *Phrasal Movement and Its Kin*. The MIT Press. ISBN: 9780262281461. DOI: <https://doi.org/10.7551/mitpress/5365.001.0001>.
- Renans, Agata, Malte Zimmermann, and Markus Greif (2011). *Questionnaire on Focus Semantics. - 2nd edition*. Working Papers of the SFB632, Interdisciplinary Studies on Information Structure (ISIS) 15. Potsdam: Universitätsverlag Potsdam.
- Rooth, Mats (1992). “A Theory of Focus Interpretation”. In: *Natural Language Semantics* 1.1, pp. 75–116. DOI: <https://doi.org/10.1007/BF02342617>.
- Salminen, Tapani (1998). “Nenets”. In: *The Uralic Languages*. Ed. by Daniel Abondolo, pp. 516–547.
- Tereschenko, Natalia M. (1989). *Slovar’ Nenezko-Russkij i Russko-Nenezkij*. Prosvezhenije.
- Wold, Dag E. (1996). “Long Distance Selective Binding: The Case of Focus”. In: *Semantics and Linguistic Theory*. Vol. 6, pp. 311–328.

Towards the semantics of Atayal polar question particles and a semantic typology¹

Sihwei Chen – Academia Sinica

Abstract. Polar questions are often associated with a bias related to the epistemic state of the speaker. This study examines the semantics of two sentence-final particles in Atayal (Formosan, Austronesian), *rwa* and *pi*, which are used to form biased polar questions. We propose that *rwa* functions as a high-level interrogative speech act applied exclusively to declaratives, while *pi* does not change the truth-conditional meaning of a polar question, but conveys the speaker’s belief that the proposition in the question form should not be included in the Common Ground. Both particles vary in certain parameters from what has been reported in the literature.

1. Introduction

1.1. Overview of the Proposal

Polar questions (PQs) in languages often carry a bias tied to the speaker’s attitude towards the question and prospective answers. This nuance, not fully captured by truth-conditional semantics for PQs, has sparked debates on the origins of this bias, the mechanisms at play, and the level at which it operates (e.g., Ladd 1981; Romero & Han 2004; a.o.). By manipulating two types of evidence – the speaker’s belief and contextual evidence (Büring & Gunlogson 2000) – this paper examines the felicity conditions of *rwa* and *pi* PQs in Atayal, contrasting them with neutral ones using *ga*. Our argument centers on two crucial dimensions: (i) whether the speaker seeks confirmation of the truth of the prejacent *p* (i.e., speaker-oriented) or confirmation that the addressee agrees that *p* is in the Common Ground (i.e., addressee-oriented) (Wiltschko 2021), and (ii) the speaker’s positive vs. negative attitude toward the truth of *p*. Specifically, we propose that *rwa* functions as a higher-level interrogative speech act applied exclusively to declaratives, while *pi* encodes a non-at-issue content that *p* should not be included in the Common Ground (cf. Frana & Rawlins 2019, originally a *FALSUM* operator in Repp 2013).

Moreover, we include a cross-linguistic comparison. This comparison reveals that higher-level PQs typically manifest a positive bias, permitting variability in their complement selection. In contrast, speaker-oriented biased PQs diverge along at least four parameters: (i) whether the bias pertains to at-issueness; (ii) whether the bias remains consistently anchored in the speaker across speech acts (i.e., anchor shift); (iii) whether the question form is affirmative or negative (i.e., polarity); (iv) whether the polarity of the question form is consistent with or opposite to the speaker’s bias (i.e., bias reversal). The first two parameters seem to vary in tandem, while all

¹ We would like to thank the audience of TripleA10 for their valuable comments. This work was also presented at the Yunshan Linguistic Salon hosted by National Taiwan Normal University, Taiwan, on Oct. 15, 2022 (via Google Meet), and at Linguistics Colloquium at Academia Sinica, Taiwan, on Oct. 16, 2023. We thank the audience at these events too. Remaining errors are our own.

possible combinations are observed for the latter, with the Atayal *pi* representing an unreported subtype.

1.2. The Atayal Language and Methodology

Atayal is an Austronesian language spoken in the northern parts of Taiwan and is considered to belong to a primary branch of Proto-Austronesian (Atayalic). It follows a predicate-initial word order, with the particles in question all in the rightmost position, referred to as sentence-final particles (SFPs):²

- (1) *cyux* *m-qwalax* *kya* (*ga* / *rwa* / *pi*)?
 PROG.DIST AV-rain there GA / RWA / PI
 'Is it raining?'

Similar to other ‘Philippine-type’ languages, Atayal features a productive and unique verbal category called voice. This category indicates the macro-semantic role of the subject through Actor Voice, Patient Voice, Locative Voice and Circumstantial Voice. Additionally, voice marking is simultaneously linked to one of the three series closely associated with mood, as depicted in Table 1.

Table 1. Atayal Voice & Mood System

	Series I: Indicative	Series II: Dependent	Series III: Hortative
Actor Voice (AV)	<i>m-/m-/Ø</i>	<i>root</i>	<i>(m-...)-a</i>
Patient Voice (PV)	<i>-un</i>	<i>-i</i>	<i>-aw</i>
Locative Voice (LV)	<i>-an</i>	<i>-i</i>	<i>-ay</i>
Circumstantial Voice (CV)	<i>s-</i>	<i>-an(i); an(i)... s-</i>	<i>-anay; anay... s-</i>

Our data is based on the Squliq dialect of Atayal spoken in Hsinchu County, Taiwan (hereafter referred to as Atayal for short). The three speakers we worked with are a woman over 65 years old, born in Jianshi, Hsinchu and married to Taoshan, Hsinchu, and a man over 74 years old, born and raised in the same village. Additionally, there is a woman around 53 years old from Jianshi, Hsinchu. However, the use of target PQ particles by this youngest woman significantly differs from the other two elder speakers. Therefore, we exclude her data from this paper. The age-related variation is an aspect worthy of further investigation.

Multiple methods are utilized in this project: (i) observing spontaneously produced textual data (which are transcribed and translated into Mandarin or English by native speakers or linguists), (ii) translation tasks, (iii) one-to-one elicitation involving controlled discourse contexts, presented either verbally or using conversational pictures (Matthewson 2004), and (iv) storyboard elicitation (as detailed in Burton and Matthewson 2015). We regard the felicity of

² Abbreviations that are not in the Leipzig Glossing Rules: AV, actor voice; COS, change of state; CV, circumstantial voice; DEP, dependent; HORT, hortative; INT, interjective; LV, locative voice; PV, patient voice.

target PQ particles within constructed contexts as the most reliable evidence for comprehending their meaning.³

2. Empirical and Theoretical Background

2.1. The Landscape of Atayal SFPs

Prior to two recent articles on different dialects, Chang (2019) and Huang (2022), the descriptions of Atayal SFPs were limited (e.g., Egerod 1999). Chang (2019) offers a comparative perspective on the SFP *ay* in C'uli' Atayal, whose occurrence is restricted to declaratives. On the other hand, Huang (2022) examines the function of each SFP through Mandarin translation and its compatibility with various clause types in Sqliq Atayal. Neither article discusses discourse contexts for SFP usage. In this subsection, we outline the distribution of the nine most common SFPs in this study and the criteria for identifying those used in polar questions. Sections 3 to 5 concentrate on the use of PQ particles.

Voice marking aids in distinguishing between declaratives, imperatives, and hortatives, specifically Series I (unmarked in gloss), II and III (glossed as 'DEP' and 'HORT', respectively) in Table 1, the first two exemplified in (2). Sentence intonations, discussed in section 2.2, work in tandem to discriminate *yes/no*-interrogatives from *wh*-interrogatives. Table 2 outlines the identifier for each clause type, along with its associated conventional speech act and presumed theoretical component.

- (2) a. hazi' wal gal-un ni Ciwas.
EPIST.POS PRF take-PV GEN Ciwas
 'Ciwas probably took it.'
- b. laxi gal-i!
NEG.IMP take-PV.DEP
 'Don't take it.'

Table 2. Basic Clause Types and Speech Acts in Atayal

Clause type (identifier)	Illocutionary force	Discourse component
declarative (final fall)	asserting	updating CG (Stalnaker 1978)
<i>wh</i> -interrogative (final fall), <i>yes/no</i> -interrogative (final rise)	asking	updating QUD (Questions Under Discussion, Roberts 1996/2012)
imperative (voice)	requesting	updating To-do List (Portner 2004)
hortative (voice)	requesting?	?

The results, classifying the 9 SFPs into two main categories – questions and assertions, are presented in Table 3 based on compatibility with five different clause types. The former are universally compatible with *yes/no*-interrogatives, whereas the latter are exclusive to assertions,

³ In the upcoming examples, contexts provided by speakers are indicated as 'offered context'.

not questions. Within each class, distinctions emerge. The SFP *ma'* forms group A2 exclusively, as it reflects a deontic modality comparable to expressions like ‘OK? alright?’, which is absent in the other members. In group A1, *ga* and *rwa* behave in the same way, which sets them apart from *pi*; moreover, only *pi* can be used outside of *yes/no*-interrogatives. Groups B, C, and D display greater diversity, and differentiation beyond clause types is required, a topic we defer for future research.

Table 3. (In)compatibility of Atayal SFPs and Clause Types

	Question				Assertion				
	A1			A2	B	C			D
	<i>ga</i>	<i>rwa</i>	<i>pi</i>	<i>ma'</i>	<i>wah</i>	<i>gaw</i>	<i>ay</i>	<i>ki</i>	<i>ru</i>
declarative	x	x	v	x	v	v	v	v	v
yes/no-interrogative	v	v	v	v	x	x	x	x	x
<i>wh</i> -interrogative	x	x	v	?	v	?	?	x	x
imperative	x	x	v	v	v	v	v	v	?
hortative	?	?	v	v	v	?	v	v	v

‘?’ means that data is yet unavailable or unclear.

2.2. PQ Particles and Sentence Intonations

Each of the three PQ particles is associated with a different intonation pattern. In the following, we first present intonation patterns without particles for comparison. Figures 1 to 3 illustrate the intonation of declaratives, *wh*-questions and *yes/no*-questions. *Yes/no*-questions are characterized by a final rise on the last syllable, which distinguishes them from the other two clause types.⁴

Figure 1. Intonation of Declarative (*mqnzyat balay qu seto' qani*. 'This student is very diligent.')

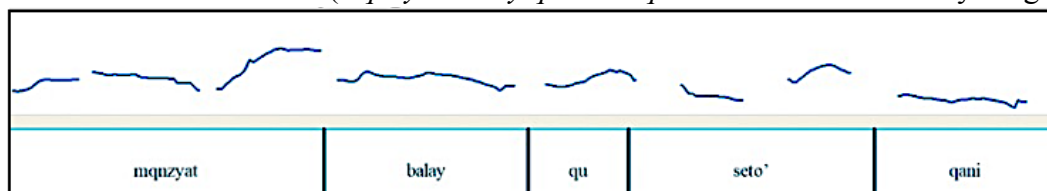
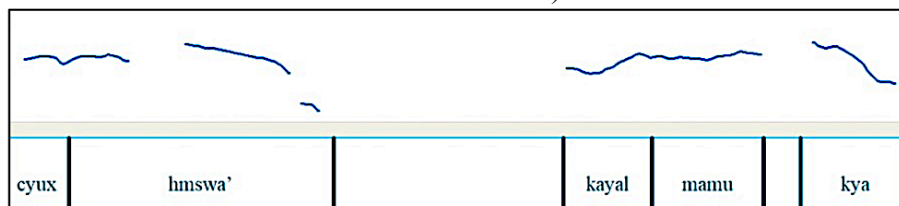
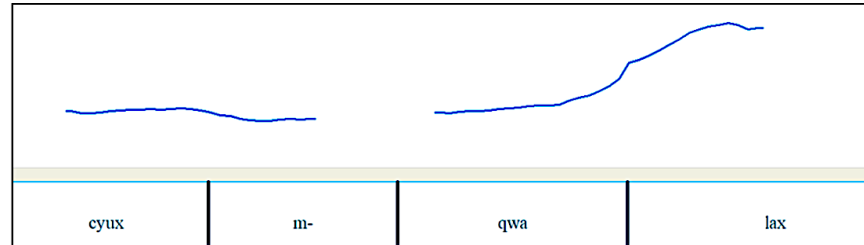


Figure 2. Intonation of *Wh*-question (*cyux hmswa' kayal mamu kya?* 'How is the weather there?')



⁴ The audio files corresponding to these spectrums are taken from <https://web.klokah.tw>.

Figure 3. Intonation of *yes/no*-question (*cyux mqwalax?* 'Is it raining?')



Figures 4 to 6 show the three patterns for PQs with particles, all of which use the question form in (1). *Ga* seems to be attached after the final rising intonation of *yes/no*-questions. With *rwa*, intonation descends throughout, resembling declaratives and *wh*-questions, possibly with a slight rise on *rwa*. Finally, *pi* behaves as if it is within the boundary of *yes/no*-questions (i.e., it carries the final rise).

Figure 4. Intonation of *ga* PQ (*cyux mqwalax ga?* 'Is it raining?')

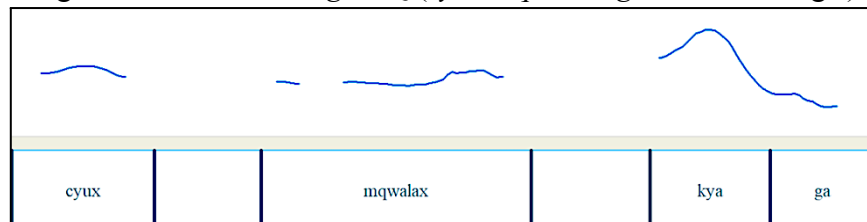


Figure 5. Intonation of *rwa* PQ (*cyux mqwalax rwa?* 'Is it raining?')

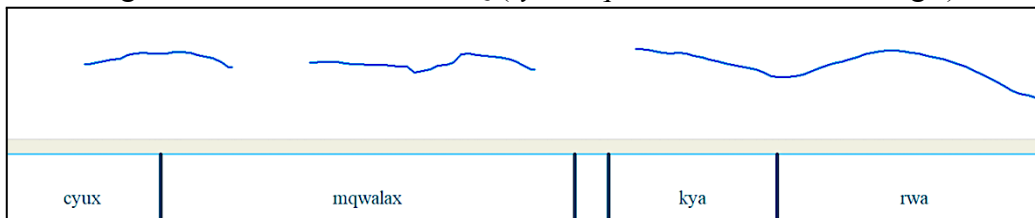
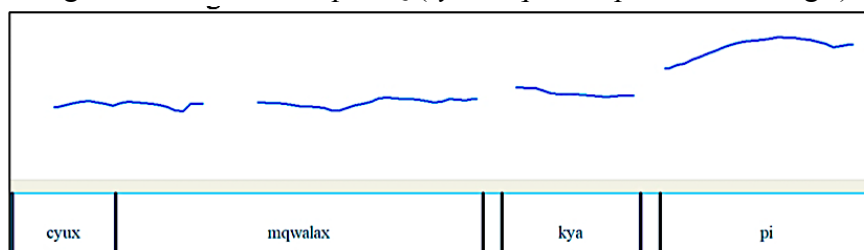


Figure 6. Intonation of *pi* PQ (*cyux mqwalax pi?* 'Is it raining?')



2.3. Perspectives: Speaker vs. Addressee

The Common Ground comprises not only shared beliefs of the interlocutors, but also their publicly displayed attitudes towards these propositions. As for the formulation of attitudes in PQs, as Dayal (2023:46) notes, “[t]hat the left periphery of interrogative clauses should have an articulated structure is hardly novel[; w]hat is novel is an explicit proposal about how that articulated structure relates to meaning.” Theories on PQs, while studying different phenomena, all introduce a distinction between speaker and addressee perspectives in the left periphery. Wiltschko (2021) proposes that confirmational markers can seek confirmation of the truth of p (Ground_{Spkr}, e.g., English *huh*) or inquire if the addressee knows p (Ground_{Adr}, e.g., English *eh*) (see section 4). Dayal (2023) compares interrogative modifiers with the phenomenon of embedded inversion in English and suggests a similar distinction. The Hindi-Urdu particle *kya*: is confined to matrix PQs and embedded clauses that are exclusively selected by predicates with interrogative complements (e.g., [+WH]); it is labelled as Perspective Phrase, addressing whether the speaker’s centered question (CQ) is active. In contrast, *Quick/Quickly* is only used in matrix questions like *Quick/Quickly, where did you hide the matza?* which require an immediate answer from the addressee and are identified as a Speech Act Phrase. Repp (2013) uses ‘CG management operators’ to refer to those operators that are not illocutionary, such as question or assertion operators, but which inform interlocutors of the CG status of a proposition, in order to resolve an epistemic conflict, e.g., VERUM/FALSUM (see section 5.2). CG-management operators are argued to take scope under illocutionary operators (Repp 2006). Figure 7 summarizes the two-way perspectives in these three theories, serving as a foundation for distinguishing the Atayal PQ particles.

Figure 7. A Comparison of Theories for the Left Periphery

[_{GroundP} Ground _{Adr}	[_{GroundP} Ground _{Spkr}	[_{CP} ...	(Wiltschko 2021)
[_{SAP} SA _{ASK}	[_{PerspectiveP} PersP _{CO}	[_{CP} ...	(Dayal 2023)
[illocutionary OP	[CG-management OP	[proposition	(Repp 2013)
		...	
speech act	perspective	clause	
		typing	

2.4. Bias in Questions

Büring & Gunlogson (2000) and Sudo (2013) propose two types of evidence for understanding bias in PQs, which we refer to as speaker belief (SB) and contextual evidence (CE). SB encompasses either the speaker’s belief or common knowledge, while CE involves contextual information available to the interlocutors in the ongoing discourse – the distinction is whether the evidence is private to the speaker or shared among interlocutors. Subsequent literature explores components of these two types of evidence and suggests decomposition. However, for our

specific purpose, we manipulate SB and CE to establish the felicity conditions of Atayal PQs. In section 4.1, we show in Atayal that CE sometimes relies on SB, aligning with the argument in Northrup (2014). Therefore, a more relevant distinction should focus on whether the speaker's epistemic evidence precedes the current CG.

Another significant question is whether negative PQs can serve as neutral questions. Despite the prevalence of affirmative PQs compared to negative ones, van Rooy and Šafářová (2003) argue that negative PQs can be neutral if the truth of the negative proposition aligns with the speaker's conversational goal. For example, in (3), learning $\neg p$ is more relevant for the speaker – specifically, to get the user to click on a button – not necessarily because the speaker has evidence for $\neg p$. Therefore, we adopt the perspective that any PQs, in principle, can be neutral or biased.

- (3) Context: When browsing the internet, a user comes across web sites asking:
Have you not been able to receive credit from your financial institution to back up your business activities? (van Rooy and Šafářová 2003:298)

3. The Neutral PQ Particle *ga*

The SFP *ga* was previously identified as the counterpart of rising intonation to mark PQs (Huang & Wu 2018:154). A PQ with rising intonation alone or with *ga* is accepted in neutral contexts:

- (4) Context: A woman is doing a survey about how many dogs there are in her neighborhood. She walks on the street, meets a man and asks:
cyux=su' q<m>azyat huzil (ga)?
 PROG.DIST=2SG.ABS raise<AV> dog GA
 'Are you raising a dog?'

The judgment for *ga* in contexts with only positive CE or SB reveals a contrast. In (5), *ga* is accepted, but in (6), the speaker's comment clearly deviates from the constructed context.

- (5) Context: I cannot see the weather outside from my bedroom. When I get up at noon today and go into the living room, I see a wet umbrella. I ask my mother:
aya', m-<in>qwalax sawni' ga?
 mom AV-<PST>rain early.today GA
 'Mom, did it rain earlier?'
- (6) Context: I heard Watan has a dog. When I meet him at the office today, I ask him:
 # *cyux=su' q<m>azyat qutux huzil ga?*
 PROG.DIST=2SG.ABS raise<AV> one dog GA
 'Do you raise a dog?' /Comment: I do not know and I have never heard of it.

We use *ga* as a standard for comparing with the other two PQ particles in Atayal. However, *ga* has specific properties worth noting. First, it is restricted to PQs, not used in declaratives,

wh-questions or alternative questions.⁵ Additionally, *ga* is not embeddable; it is always fixed to the addressee if it encodes a PersP_{CO}, unlike Hindi-Urdu *kya*:. This hypothesis fits with its intonation pattern (Figure 4), resembling the biased *rwa*, but not *pi* (sections 4–5).

4. The Addressee-oriented, Positively-biased PQ particle *rwa*

4.1. Bias Profile

Unlike *ga*, PQs with *rwa* require a positive SB. The contexts in (7)–(8), both of which contain no prior belief in *p*, are neutral, and in these contexts PQs with *rwa* are rejected (while PQs with *ga* are given voluntarily). Note that the difference in the polarity of these questions is due to the speaker's different conversational goals: to find out about the weather in (7) and to attract the attention of parents whose children have learning difficulties in (8).

- (7) Context: I call my friend in Peru from Taipei. I ask, “What's the weather like over there? _____”
- | | | | | |
|---------------|--|-----------------|------------|-------------|
| # <i>cyux</i> | | <i>m-qwalax</i> | <i>kya</i> | <i>rwa?</i> |
| PROG.DIST | | AV-rain | there | RWA |
- 'Is it raining there?' /Comment: I am sure it is raining over there.
- (8) Context: I am walking down the street and see a billboard that says, “_____ Try our medicine.”
- | | | | | | |
|---------------|----------------|---------------|---------------|------------------|-------------|
| # <i>ini'</i> | <i>thuzyay</i> | <i>inblaq</i> | <i>m-qbaq</i> | <i>laqi'=su'</i> | <i>rwa?</i> |
| NEG | able.AV.DEP | well | AV-study | child=2SG.GEN | RWA |
- 'Your child is not able to study well, right?' /Comment: I know the situation of his child. We are very close. The ads on the street will not use it.

In contrast, *rwa* is accepted in contexts where the speaker has gained private evidence for *p*:

- (9) Context: the same as in (6)
- | | | | | |
|-------------------|------------------------|--------------|--------------|-------------|
| <i>cyux=su'</i> | <i>q<m>azyat</i> | <i>huzil</i> | <i>balay</i> | <i>rwa?</i> |
| PROG.DIST=2SG.ABS | raise<AV> | dog | really | RWA |
- 'You are really raising a dog, right?' /Comment: I've heard this before. It's a question with a positive bias.

However, in contexts with exclusively positive CE, (10)–(11), the judgement for *rwa* is less straightforward. Substituting the volunteered *ga* PQs with *rwa*, the speaker emphasizes stronger positive epistemic evidence, therefore marked as %. Assuming with Northrup (2014) that positive CE can sometimes become part of the speaker's private belief (but not vice versa), we argue that contexts like (10)–(11) are a subtype of those with a positive SB. Contrasting (7)–(8) with (9)–(11) suggests that *rwa* encodes a positive epistemic stance of the speaker. Table 4 summarizes the felicity differences of *ga* and *rwa*.

⁵ Huang (2022:280–281) reports examples of *ga* in imperatives and declaratives in the Jianshi dialect, but my speaker consistently rejects such uses. This needs to be investigated further.

- (10) Context: I've only recently met Tali'. Today I go to the library with him. I notice that he borrows many books about the history of Taiwan. I ask him:
 % *s<m>oya'=su' mita' biru' na rekisi rwa?*
 like<AV>=2SG.ABS see.AV book GEN history RWA
 'You like reading history books, don't you?' /Comment: If he borrows **a lot of them and not just one or two** [emphasis added], I know he likes them.
- (11) Context: Tali' is my new friend. Today I have dinner with him. He orders all kinds of meat but no beef. I ask him:
 % *ini'=su' soya' maniq hi' na kacing rwa?*
 NEG=2SG.ABS like.AV.DEP eat.AV meat GEN bull RWA
 'Do you not like to eat beef?' /Comment: **I am sure** [emphasis added] he does not eat beef, so he does not order it.

Table 4. The Felicity Conditions of *ga* and *rwa*

	Neutral	Positive SB	Positive CE	Bias
<i>ga</i>	✓ (4)	# (6)	✓ (5)	neutral
<i>rwa</i>	# (7), (8)	✓ (9)	% (10), (11)	positive SB

That *rwa* requires a positive epistemic inference on the part of the speaker is supported by its use as paraphrase of 'High Negation PQs,' even when the PQs lack negation. HiNegPQs refer to PQs with preposed negation in English (Ladd 1981; Romero & Han 2004). For example, *Isn't Jane coming too?* carries a positive epistemic inference despite the negative question form.⁶ Romero & Han (2004) analyze this inference as a pragmatic implicature, explaining that the speaker is essentially asking, "Are you not 100% sure that *p*?", from which the hearer infers that the speaker has previously endorsed *p*. (12)–(13) are situations where HiNegPQs are felicitous, characterized by opposing beliefs between the speaker and the hearer. In these contexts, Atayal *rwa* is volunteered along with the sentence-initial particle *aw'*:⁷

- (12) Context: I invite Tali' for a drink and tell him to come after dinner. When he arrives, he asks if I have anything to eat. I say: (adopted from Frana & Rawlins 2019:16)
iyat=su' m-<in>aniq nga? aw'=su' m-<n>aniq lrwa?
 NEG=2SG.ABS AV-<PST>eat first.GA? yes=2SG.ABS AV.eat<PST> COS.RWA
 'Did you not eat? You have eaten, haven't you?' /Comment: The two sentences are both uttered by you; you know he has eaten and simply answer yourself.
- (13) Context: Rimuy invites us for dinner. At dinner Tali' barely touches any food. Rimuy asks him: (adopted from Frana & Rawlins 2019:16)

⁶ Romero & Han (2004) argue that HiNegPQs are ambiguous depending on whether an epistemic operator scopes over or under the negation (outer vs. inner reading, cf. section 5.2) but the ambiguity is controversial (cf. Goodhue 2022). In contrast, PQs with negation in-situ are termed LoNegPQs (e.g., *Is Jane not coming?*).

⁷ The particle *aw'*, literally 'yes', signals a contrast of expectation in the discourse and is not specific to the use of *rwa'*.

aw'=su' ini' qaniq nrwa? wa cikuy niq-un=su'.
 yes=2SG.ABS NEG eat.AV.DEP first.RWA INT little eat-PV=2SG.ERG
 'You didn't eat first, did you?' 'What you ate is little.'

4.2. Evidence for Encoding Addressee's Perspective

Given the generalization that *rwa* implies a positive bias, the question arises: what does the speaker want to confirm? If it is the truth of *p*, *rwa* PQs are speaker-oriented. If it is whether *p* is also in their interlocutor's belief, *rwa* PQs are addressee-oriented. There is also a possibility that *rwa* PQs are ambiguous. We provide evidence supporting the unambiguous addressee-oriented interpretation.

First, addressee-oriented biased PQs, unlike speaker-oriented ones, do not need confirmation of the truth of *p*. They are expected to be felicitous when the speaker believes in or commits to *p*. This is evident in (14), where it is obvious to the interlocutors that Tali' got surprised, so it is unnatural for the speaker to require confirmation of this.

- (14) Context: Rimuy has organized a surprise party for Tali'. As Tali' enters the room, everyone shouts "surprise." Observing his surprised expression, Rimuy utters: (adopted from Wiltschko 2021:113)
- m-nkux=su' lrwa?*
 AV-get.scared=2SG.ABS COS.RWA
 'You got scared, right?' /Comment: You asked that after he opened the door and got surprised.

The situation is similar when *p* is an evaluative judgment, such as with predicates of personal taste. As far as perspective is concerned, the truth of *p* is relative to the speaker in assertions and relative to the addressee in questions (referred to as 'interrogative flip'). As shown in (15), the speaker can felicitously use *rwa* to ask whether the addressee agrees with his evaluation, akin to the Canadian *eh*. The behavior of Atayal *rwa* contrasts with the English *huh*, which is bad in (15), as *huh* would suggest that the speaker is asking for confirmation of correctness of his own taste.

- (15) Context: A and B go to the movies and both enjoy it. As they leave, they talk about the movie and A says: (adopted from Wiltschko 2021:122)
- a. *blaq ngayan ega' qani rwa?*
 good.AV watch.LV movie this RWA
 'This is a good movie, right?' /Comment: I like it myself and I think you like it too.
- b. *This was such a good movie, {eh/#huh}?*

The difference in authority between interlocutors also serves as a diagnosis. If the addressee holds authority over the speaker, it is less appropriate to seek the speaker's agreement, as shown by the fact that *rwa* is infelicitous in (16); instead, affirmation of the truth of *p* is more appropriate. Conversely, when the speaker has absolute authority, especially in personal matters, the use of *rwa* is still possible to ask for the confirmation of others, as in (17) (while a speaker-oriented one like the English *huh* is not possible, e.g., *??I have a dog, huh?*).

- (16) Context: An employee converses with his boss after the employee has made an unforgivable mistake. Employee to boss: "I am fired, huh?" (adopted from Wiltschko 2021:121)

lax-an=saku' *lrwa?*
 abandon-LV=1SG.ABS COS.RWA

'You are going to abandon me, aren't you?' /Comment: I would say this if I realized that the boss was only asking other people, but not me, to do the work. I would not say this immediately if I made the mistake.

- (17) Offered context: A and B are close. They have a conversation with a new friend C. A says to C: "I have a dog." Then A turns to B:

nyux=saku' *q<m>azyat* *huzil* *rwa?*
 PROG.DIST=1SG.ABS raise<AV> dog RWA

'I am raising a dog, right?' /Comment: Help me confirm with C that I have a dog.

Since *rwa* is only felicitous in situations where the speaker seeks the addressee's agreement, it is unambiguously oriented towards the addressee.

4.3. Towards a Formalization and Parameters of Variation in high-level questions

A notable characteristic of Canadian *eh* is its flexibility with different clause types, as long as it is used to assume that the interlocutor agrees with the speech act, e.g., *What are they trying to do, eh?* *Think about it, eh?* and *What a game, eh?* (Wiltschko 2021:108; 110; 114). Another comparable instance is Cantonese *ho2*, supported by its ability to stack on other question SFPs (Lam 2014; Law et al. 2018). In contrast, Atayal *rwa* exclusively attaches to a declarative host; for instance, (18). The SFP *honnh* in Taiwanese Southern Min shares a similar limitation (Lâu 2022).

- (18) Context: Tali' and Rimuy are in a public lecture. The lecture is rather obscure and Tali' does not follow it. From Rimuy's face, he determines that she doesn't understand the lecture either. Tali' says: (adopted from Wiltschko 2021:107)

nanu' *nyux=nya'* *s-kayal* *rwa?*
 what PROG.PROX=3SG.ERG CV-speak RWA

Intended: 'What is he talking about? Do you share the same question?'

To account for the use of *rwa*, we modify the analysis of Cantonese *ho2* by Law et al. (2018) by adding a selection restriction on the complement. Before presenting the formula, the notations are as follows (cf. Law et al. 2018:60ff.). A context *c* constitutes a pair of $\langle cs_c, J_c \rangle$, where cs_c is

the context set $\subseteq W$ (roughly, the CG) and J_c is a stack of salient propositions in the CG. Making an assertion means pushing a proposal onto the stack for the addressee to confirm, represented as $push(e, J_c)$, while asking a polar question minimally differs in pushing a set of two potential updates. Moreover, speech acts themselves are operable (e.g., Krifka 2015). Addressee-oriented confirmationals literally ask whether the addressee can felicitously perform the same speech act as the speaker; this is indicated by the subscripts s_c and a_c , referring to the speaker/addressee of c . As in (19)a, Atayal *rwa* turns an assertion into a new interrogative speech act asking whether the addressee also asserts it. The example in (17)b is computed as in (19)b.

$$(19) \quad a. \quad \llbracket rwa \rrbracket_{w, s_c, a_c}(p) = rwa(assert(p))_{w, s_c, a_c} =$$

$$\langle cs_c, push\left(\left\{cs_c \cap \left\{w' \mid assert_{w, a_c} \cdot (p) \text{ is defined in } w'\right\}, cs_c \cap \left\{w' \mid assert_{w, a_c} \cdot (p) \text{ is undefined in } w'\right\}\right), J_c\right\rangle$$

$$b. \quad (17)b =$$

$$\langle cs_c, push\left(\left\{cs_c \cap \left\{w' \mid assert_{w, a_c} \cdot (\llbracket nyux saku' qmazyat huzil \rrbracket) \text{ is defined in } w'\right\}, cs_c \cap \left\{w' \mid assert_{w, a_c} \cdot (\llbracket nyux saku' qmazyat huzil \rrbracket) \text{ is undefined in } w'\right\}\right), J_c\right\rangle$$

defined only if $assert_{w, s_c} \cdot (\llbracket nyux saku' qmazyat huzil \rrbracket)$ is defined in w

Table 5 compares the four addressee-oriented biased PQ particles discussed above. The restriction on the complement appears to be an important parameter on which languages vary. Atayal *rwa* can also be used in narratives like *eh*, which we do not discuss for reasons of space (see e.g., Rau 1992:247), but this feature seems to hold in the four languages.

Table 5. Variation in Addressee-oriented Biased PQs

	Ambiguous	Restriction on complement	Narrative use
Canadian <i>eh</i>	✓	✗	✓
Cantonese <i>ho2</i>	✗	✗	?
TSM <i>honnh</i>	✗	✓	✓
Atayal <i>rwa</i>	✗	✓	✓

5. The Speaker-oriented, Non-positively-biased PQ particle *pi*

5.1. Bias Profile

Another biased PQ particle in Atayal is *pi*, as shown by its infelicity in neutral contexts:

- (20) Context: A woman is doing a survey about how many dogs there are in her neighborhood. She meets a man on the street and asks him:
- # *cyux=su'* *q<m>azyat huzil pi?*
 PROG.DIST=2SG.ABS raise<AV> dog PI
 'Do you have a dog?' /Comment: You really need to tell me. Quite rude and dubious: "Do you really have a dog and not the other animal?"

- (21) Context: I walk past a Chinese clinic and see the advertisement: "Do you have trouble sleeping? _____. Chinese medicine helps you sleep better without side effects."
 # *ini'=su'* *sawya'* *maniq* *iyu'* *ka* *inblaq* *m-'abi'* *pi?*
 NEG=2SG.ABS like.AV.DEP eat.AV medicine REL well AV-sleep PI
 'Do you not like taking medicine that helps you sleep?' /Comment: You actually want to have it.'

The most common contexts for *pi* involve conflicting evidence: the speaker may be a priori in favor of $\neg p$, but context suggests *p* (22), or vice versa (23). Notice that *pi* differs from the biased particle *rwa* in that it has a question form whose polarity is opposed to the speaker's prior epistemic bias and aligns with CE instead. We refer to this as 'bias reversal' (cf. Han 2002).

- (22) Offered context [SB: $\neg p$, CE: *p*]: When I got up, I saw that it was dry on the floor. My mother says it was raining. I ask her:
aya', *m-<in>qwalax* *sawni'* *pi?* [p *pi*]
 Mom AV-<PST>rain early.today PI
 'Mom, did it rain earlier?'
- (23) Context [SB: *p*, CE: $\neg p$]: I remember that Watan likes to eat meat, including beef. Today, I have dinner with him. He orders all kinds of meat except beef. I ask:
ini'=su' *soya'* *maniq* *hi'* *na* *kacing* *pi?* [$\neg p$ *pi*]
 NEG=2SG.ABS like.AV.DEP eat.AV meat GEN bull PI
 'Don't you like eating beef?' /Comment: I'm sure that you like to eat beef.

PQs exhibiting bias reversal are commonly called negatively biased PQs, seen in languages like Cantonese *mei* (Lam 2014) and Medumba *-á* (Keupdjio & Wiltschko 2018). The question arises as to whether the Atayal *pi* belongs to this category – whether it **requires** conflicting evidence. The following examples suggest a negative answer, since none of them contains contradictory beliefs between the interlocutors.

- (24) Context: Two girls gossip about whether Tali' has a girlfriend. A says, "Maybe Tali' has a girlfriend." B says, "Maybe he does not have one." They ask me:
hazi' *kya* *rangi'* *kneril* *qu* *Tali'* {*pi*/*rwa*/*ga*}?
 perhaps exist friend woman ABS Tali' PI/RWA/GA
 'Does Tali' perhaps have a girlfriend?' /Comment: You guys do not know so ask other people.
- (25) Offered context: My friend recommends a movie I haven't watched yet. I ask him:
blaq *ngayan* *ega'* *qani* {*pi*/*rwa*/*ga*}?
 good.AV watch.LV movie this PI/RWA/GA
 'Is the movie good?' /Comment: You confirm with him because you are going to watch it.

- (26) Offered context: I want to borrow some books to read in my spare time, but I don't know which books to borrow. My friend suggests:

s<m>oya'=su' mita' biru' na rekisi {pi/#rwa/#ga}?
 like<AV>=2SG.ABS see.AV book GEN history PI/RWA/GA
 'Do you like to read history books?'

We argue that what these contexts have in common is the speaker's uncertainty about whether p or $\neg p$ is true. In (24), both A and B lack concrete evidence for Tali's relationship status and seek an answer from others; in (25)–(26), the speaker may hesitate to accept p or doubt that it is true. Neither *rwa* nor *ga* is accepted in these contexts; the infelicity of *ga* indicates that the contexts are not simply neutral. The felicity conditions of *pi* are summarized in Table 6.

Building on (i) the bias reversal and (ii) the generalization that the speaker is not sure of the proposition conveyed by the question form, we propose that *pi* encodes a FALSUM operator as a non-at-issue content (cf. Potts 2005). Before presenting the proposal, we discuss the background of FALSUM (section 5.2) and a minimally different PQ particle in Italian (section 5.3).

Table 6. The Felicity Conditions of *pi*

SB	CE	Intent	Comparison
$\neg p$	p	check p	<i>really p</i>
p	$\neg p$	check $\neg p$	<i>really \neg p</i> ; HiNegPQ _{inner} (cf. footnote 6)
$p \vee \neg p$	$p \vee \neg p$	check p ⁸	emphatic <i>do</i>

5.2. VERUM and FALSUM

The analysis of English HiNegPQs by Romero & Han (2004) proposes a VERUM operator under negation, the interaction of which yields the intended positive bias, (27). VERUM is a conversational epistemic modal (or a CG-management operator, section 2.3). (28) says that in all worlds w' that conform to x 's knowledge in w that in all the worlds w'' satisfying x 's conversational goals in w' , p is added to the CG. Note that the reference of x in questions in English is shifted from the speaker to the addressee, while it remains the same for particles such as Italian *mica* and Atayal *pi* (section 5.3).

- (27) Isn't Jane coming too?
 [_{CP} Q **not** [**VERUM** [_{IP} Jane is coming] too]]
 ≈ Are you not 100% sure that we should add to CG that p ?

⁸ I assume that negation is a marked form and can only be felicitously used when the positive counterpart was previously taken to be most likely true (cf. van Rooy and Šafářová 2003).

$$(28) \quad \llbracket \text{VERUM}_i \rrbracket^{c[x/i]} = \lambda p_{\langle s, t \rangle}. \lambda w. \forall w' \in \text{Epi}_x(w) [\forall w'' \in \text{Conv}_x(w') [p \in \text{CG}_{w''}]]$$

(Romero 2015:523)⁹

Repp (2006, 2013) includes the polar antonym of VERUM in the inventory, called FALSUM (29). While VERUM indicates that the degree of strength for adding p to the CG is 100%, for FALSUM the degree of strength for adding p to the CG is zero; in other words, p should not be added.

$$(29) \quad \llbracket \text{FALSUM}_i \rrbracket^{c[x/i]} = \lambda p_{\langle s, t \rangle}. \neg p$$

Non-at-issue: $\lambda p_{\langle s, t \rangle}. \lambda w. \forall w' \in \text{Epi}_x(w) [\forall w'' \in \text{Conv}_x(w') [p \notin \text{CG}_{w''}]]$ (Based on Romero 2015:524; cf. Repp 2013:239)

5.3. Atayal *pi*: a Variant of Italian *mica*

In Italian, the particle *mica*, grammaticalized from negation, is argued to encode FALSUM as its presupposition (Frana & Rawlins 2019). Unlike Atayal *pi*, *mica* has more restricted felicity conditions, felicitous only when the speaker expects $\neg p$ but CE suggests p . For instance, both *mica* and *pi* are accepted in (30), where Tali's behavior implies that he might have eaten, contrary to the speaker's expectation that he should not have. However, *pi* is also felicitous in contexts where the speaker expects p but CE suggests $\neg p$ (e.g., (23)) but *mica* is not (Frana & Rawlins 2019:3; 16).

- (30) Context (= (13)): Rimuy invites us for dinner. At dinner Tali' barely touches any food. Rimuy asks him: (adopted from Frana & Rawlins 2019:16)
- a. *Mica hai già mangiato?*
'You didn't eat first, did you?'
 - b. *m-<in>aniq=su' l.pi?*
AV-<PST>eat=2SG.ABS COS.PI
'You already ate, didn't you?'

The pair in (30) also highlights another difference between the two particles. The question form modified by the Atayal *pi* exhibits bias reversal, meaning it opposes to the speaker's bias. In contrast, for the Italian *mica*, the question form always aligns with the speaker's negative attitude.¹⁰ Leveraging this difference, we only have to adjust one aspect in Frana & Rawlins's proposal for *mica*. As per (31), *mica* operates by taking a proposition, providing its negation, and concurrently signaling the speaker's certainty that, within all the worlds aligning with their conversational goals, p is not in the CG. *Spkr/i* indicates that *pi* is lexically bound to the speaker.

$$(31) \quad \llbracket mica_i \rrbracket^{c[Spkr/i]} = \lambda p_{\langle s, t \rangle}. \neg p$$

⁹ $\text{Epi}_x(w)$ is the set of worlds that conform to x 's knowledge in w ; $\text{Conv}_x(w')$ is the set of worlds where all the conversational goals of x in w' are fulfilled.

¹⁰ This explains why *mica* is not used when the speaker's epistemic stance is affirmative.

Non-at-issue: $\lambda p_{\langle s,t \rangle} \cdot \lambda w. \forall w' \in \text{Epi}_{\text{Spkr}}(w) [\forall w'' \in \text{Conv}_{\text{Spkr}}(w') [p \notin \text{CG}_{w''}]]$
 (Based on Romero 2015:524; cf. Repp 2013:239)

Atayal *pi* serves as a positive counterpart of *mica* but shares with it the characteristic of being consistently anchored in the speaker, (32). The truth conditions of example (22) are given in (33): by asking in the form *p pi*, the speaker also expresses his attitude that he does not believe *p* is part of the CG. The pragmatic reasoning for opting for *pi* PQs is as follows: The speaker chooses a question form with *pi* over one with the neutral one *ga* for a reason – either due to a distinct prior belief than the addressee or because the context does not support either belief.

(32) $[[pi_i]]^{c[\text{Spkr}/i]} = \lambda p_{\langle s,t \rangle} \cdot \mathbf{p}$
 Non-at-issue: $\lambda p_{\langle s,t \rangle} \cdot \lambda w. \forall w' \in \text{Epi}_{\text{Spkr}}(w) [\forall w'' \in \text{Conv}_{\text{Spkr}}(w') [p \notin \text{CG}_{w''}]]$

(33) a. *minqwalax sawni' pi?* (= (22))
 'Did it rain earlier?'
 b. $[[[\mathbf{Q} [pi_i [minqwalax sawni']]]]]^{c[\text{Spkr}/i]} = \{p: minqwalax sawni', \neg p: iyat minqwalax sawni'\}$
 Non-at-issue: $\lambda p_{\langle s,t \rangle} \cdot \lambda w. \forall w' \in \text{Epi}_{\text{Spkr}}(w) [\forall w'' \in \text{Conv}_{\text{Spkr}}(w') [p \notin \text{CG}_{w''}]]$

A remaining issue that we cannot address here is the use of *pi* outside PQs, in declarative, *wh*-questions, imperatives, etc. (Table 3). (34) illustrates *pi* in *wh*-questions, where it gives rise to a ‘non-existence effect’ parallel to that in rhetorical questions: ‘He will not take us to any restaurant.’ We suspect that the reversal bias parameter is at play here, canceling out the existential presupposition in *wh*- questions. This aspect has to be left to another occasion.

(34) Offered context: My brother wants to take us to dinner, but we are just walking down the street. I have no idea what restaurant he's going to take us to:
ras-un=ta'=nya' maniq inu' pi?
 take-PV=1PL.INCL=3SG.ERG eat.AV where PI
 'Where will he take us to eat?'

6. Conclusion and Implications

We showed that Atayal SFPs *rwa* and *pi* form biased PQs with two distinct characteristics. Firstly, *rwa* asks the addressee to felicitously assert *p*, functioning as a high-level interrogative speech act, while *pi* patterns like typical PQs to verify *p*. Secondly, both encode the speaker’s epistemic stance, albeit in opposing directions: *rwa* aligns the speaker’s belief with the question form, while *pi* signals to the addressee that the speaker’s belief contradicts the question form or is uncertain about it. These distinctions align with our initial examination of the prosody: *rwa* follows a falling intonation, commonly linked with declaratives (Figure 5), signifying that it selects an assertion as input. Conversely, the non-positively-biased *pi* itself is within a rising intonation, typically linked with *yes/no*-questions (Figure 6), suggesting that *pi* modifies the speaker’s attitude in questioning.

The results from Atayal bear significant implications for cross-linguistic variation. Comparing Atayal *pi* with Italian *mica* and Mandarin *ba* (Chen 2022) shows that CG-management operators can vary in at least four ways, Table 7. In particular, the polarity of the question form and the presence of bias reversal are two features that differ. Further comparison of these PQs with English HiNegPQs, LoNegPQs and Mandarin *bushi* suggests a potential link between at-issuedness and the shift in the anchor of the attitude in PQs. Additional research is needed to explore this connection.

Table 7. Potential Variation in CG-management operators in PQs

	At-issue?	Polarity?	Bias reversal?	Anchor shift?
Atayal <i>pi</i>	non-at-issued	affirmative	yes	no
Mandarin <i>ba</i> (?)	non-at-issued	affirmative	no	no
?	non-at-issued	negative	yes	no
Italian <i>mica</i>	non-at-issued	negative	no	no
?	at-issue	affirmative	yes	yes
Mandarin <i>bushi</i> (?)	at-issue	affirmative	no	yes
HiNegPQs	at-issue (?)	negative	yes	yes
LoNegPQs	at-issue	negative	no	yes

References

- Büring, Daniel & Gunlogson, Christine. 2000. Aren't positive and negative polar questions the same?. (https://semanticsarchive.net/Archive/mYwOGNhO/polar_questions.pdf) (Accessed 2023-11-28)
- Burton, Strang & Matthewson, Lisa. 2015. Targeted construction storyboards in semantic fieldwork. In Ryan M. Bochnak & Lisa Matthewson (eds.), *Methodologies in semantic fieldwork*, 135–156. New York: Oxford University Press.
- Chang, Henry Y. 2019. Formosan speech act mood in comparative syntax. In Matt Pearson (ed.), *Journal of the Southeast Asian Linguistics Society Special Publication No.4: Papers from the Austronesian Formal Linguistics Association 24*, 1–13. University of Hawai'i Press.
- Chen, Sihwei. 2022. Bias in Mandarin polar questions: An acceptability judgment study. Paper presented at the 14th Workshop on Formal Syntax & Semantics (FOSS-14), Taipei, 21–22 Oct. 2022.
- Dayal, Veneeta. 2023. The interrogative left periphery: How a clause becomes a question. *Linguistic Inquiry*. 1–85.
- Egerod, Søren. 1999. *Atayal-English dictionary*. 2nd edn. Copenhagen: The Royal Danish Academy of Sciences and Letters.

- Frana, Ilaria & Rawlins, Kyle. 2019. Attitudes in discourse: Italian polar questions and the particle mica. *Semantics & Pragmatics* 12(16). 1–48.
- Han, Chung-Hye. 2002. Interpreting interrogatives as rhetorical questions. *Lingua* 112. 201–229.
- Huang, Lillian Mei-Jin & Wu, Xin-Sheng. 2018. *Taiyayu yufa gailun* [Introduction to Atayal Grammar]. Taipei: Council of Indigenous Peoples.
- Huang, Lillian Mei-Jin. 2022. *Taiyayu cilei jiqi jiaoxue* [Part of speech in Atayal and its pedagogy]. Taipei: Indigenous Language Research and Development Foundation.
- Goodhue, Daniel. 2022. All focus is contrastive: On polarity (verum) focus, answer focus, contrastive focus and givenness. *Journal of Semantics* 39(1). 117–158.
- Keupdjio, Hermann & Wiltschko, Martina. 2018. Polar questions in Bamileke Medumba. *Journal of West African Languages* 45(2). 18–40.
- Krifka, Manfred. 2015. Bias in commitment space semantics: Declarative questions, negated questions, and question tags. *Semantics and Linguistic Theory* 25. 328–345.
- Ladd. 1981. A first look at the semantics and pragmatics of negative questions and tag questions. In Roberta A. Hendrick et al. (eds.), *Papers from the Seventeenth Regional Meeting, Chicago Linguistic Society*, 164–171. Chicago: Chicago Linguistic Society.
- Lam, Zoe Wai-Man. 2014. A complex ForceP for speaker- and addressee-oriented discourse particles in Cantonese. *Studies in Chinese Linguistics* 35(2). 61–80.
- Lâu, Sêng-hiân. 2022. Sentence-final outliers and inter-sentential syntax. Paper presented at the 14th Workshop on Formal Syntax & Semantics (FOSS-14), Taipei, 21–22 Oct. 2022.
- Law, Jess H. K. & Li, Haoze & Bhadra, Diti. 2018. Questioning speech acts. *Proceedings of Sinn und Bedeutung* 22(2). 53–70.
- Matthewson, Lisa. 2004. On the methodology of semantic fieldwork. *International Journal of American Linguistics* 70(4). 369–415.
- Northrup, Oliver. 2014. *Grounds for commitment*. Doctoral dissertation, University of California, Santa Cruz.
- Portner, Paul. 2004. The semantics of imperatives within a theory of clause types. In Robert B. Young (ed), *Proceedings of Semantics and Linguistic Theory 14*, 235–252. LSA.
- Potts, Christopher. 2005. *The logic of conventional implicatures*. Oxford: Oxford University Press.
- Rau, Der-hwa Victoria. 1992. *A grammar of Atayal*. Doctoral dissertation, Cornell University.
- Repp, Sophie. 2006. ¬(A&B). Gapping, negation and speech act operators. *Research on Language and Computation* 4(4). 397–423.
- Repp, Sophie, 2013. Common ground management: modal particles, illocutionary negation and verum. In Daniel Gutzmann & Hans-Martin Gärtner (eds.), *Beyond expressives: Explorations in use-conditional meaning*, 231–274. Leiden: Brill.
- Roberts, Craige. 2012. Information structure: Towards an integrated formal theory of pragmatics. *Semantics and Pragmatics* 5(6). 1–69. (Republish of Roberts 1996.)
- Romero, Maribel. 2015. High negation in subjunctive conditionals and polar questions. *Proceedings of Sinn und Bedeutung* 19. 519–536.
- Romero, Maribel & Han, Chung-hye. 2004. On negative yes/no questions. *Linguistics and Philosophy* 27. 609–658.
- Stalnaker, Robert C. 1978. Assertion. In Peter Cole (ed.), *Pragmatics*, 315–332. New York: Academic Press.

- Sudo, Yasutada. 2013. Biased polar questions in English and Japanese. In Daniel Gutzmann & Hans-Martin Gärtner (eds.), *Beyond expressives: Explorations in use-conditional meaning*, 275-295. Leiden: Brill.
- van Rooy, Robert & Marie Šafářová. 2003. On polar questions. In Robert B. Young & Yuping Zhou (eds.), *Proceedings of Semantics and Linguistic Theory* 13, 292–309. Fort Washington: CLC Publications.
- Wiltschko, Martina. 2021. *The grammar of interactional language*. Cambridge: Cambridge University Press.

In search of exclusive plural — insights from Hausa¹

Agata Renans — Ruhr-Universität Bochum

Abstract. There is an ongoing debate in the literature whether there is a language with the exclusive plural only, i.e., in which plural invariably gives rise to the more-than-one meaning both in upward and downward entailing contexts. I extend the discussion on the exclusive vs. inclusive plural in a novel way to the interpretation of the verbal plural, so-called pluractionals. In particular, I examine the meaning of plural and pluractionals in Hausa, a language for which it has been claimed in the previous literature that it has the exclusive plural only, both in the nominal and verbal domain (see e.g., Součková and Buba 2008, Součková 2011, Zimmermann 2021). The outcome of the empirical elicitation studies suggests that the inclusive interpretation in Hausa is available in both domains. Thus the data from Hausa inform an empirically adequate theory of plural across domains and languages.

1 Introduction

There is a vivid ongoing discussion in the plurality-literature on whether there is a language with the exclusive-only plural, i.e., a language in which plural invariably gives rise to the more-than-one meaning both in upward and downward entailing contexts. This paper aims at shedding new light on this debate by contributing novel data from the Hausa language and by extending the discussion on the exclusive-inclusive plural to the verbal domain, i.e., pluractionality – a step that to the best of my knowledge has not been taken in the theoretical literature so far.

The paper focuses on the Hausa language, as it shows interesting properties from the point of view of the discussion on the meaning of plural: First, there are overt plural and pluractional markers in Hausa, i.e., morphological plurals in the nominal and verbal domain. Second, it has been claimed in the previous literature that in Hausa there is an exclusive-only plural both in the nominal and verbal domain (Zimmermann 2008, 2021, Součková 2011).

Contrary to latter claim, I will argue in this paper that up to now there is no conclusive empirical evidence across languages for the existence of the exclusive-only plural either in the nominal or verbal domain. These observations are crucial as they constrain cross-linguistically adequate theories on the meaning of plural.

The structure of the paper is as follows. Section 2 gives an overview of the discussion on the meaning of plural, focusing on the exclusive vs. inclusive debate. Sections 3 and 4 contribute new data from Hausa to this debate, concentrating on the nominal plural in Section 3 and on the verbal plural in Section 4. Section 5 discusses an interaction of the number morphology with numerals and Section 6 summarizes.

¹I would like to thank Yusuf Baba Gari for help with the Hausa data and discussion. I would also like to thank the audience at TripleA 10 in Potsdam and Linguistischer Arbeitskreis in Köln for inspiring discussions on the material presented in this paper.

2 Meaning of Plural

2.1 Inclusive vs. exclusive plural

The initial observation in the research on the meaning of plural is that the meaning conveyed by (1) with a plural count noun is very much different from the meaning conveyed by (2) with a singular count noun:

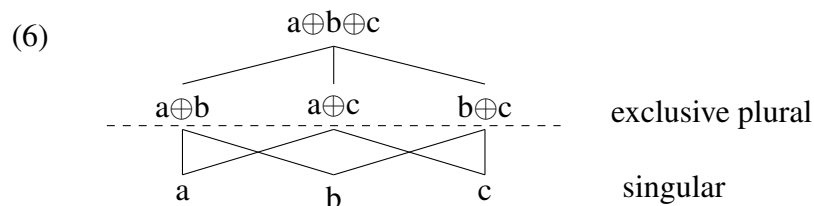
- (1) Yesterday morning, Malami bought (some) books.
 (2) Yesterday morning, Malami bought some book.

The first, intuitively compelling hypothesis is that while (1) with a plural count noun conveys the meaning that Malami bought more than one book, (2) with a singular count noun does not:

- (3) Yesterday, Malami bought (some) books.
 \rightsquigarrow *Yesterday, Malami bought more than one book*
 (4) Yesterday, Malami bought some book.
 $\not\rightsquigarrow$ *Yesterday, Malami bought more than one book*

It follows that under this first hypothesis, plural is interpreted exclusively. That is, atomic elements are excluded from the denotation of plural (see e.g., Link 1983, Chierchia 1998, Harbour 2014), as presented schematically in (5) and (6) below:

- (5) a. $[[\text{book}]] = \{a,b,c\}$ SINGULAR
 b. $[[\text{book-s}]] = \{\{a,b\}, \{b,c\}, \{a,c\}, \{a,b,c\}\}$ EXCLUSIVE PLURAL



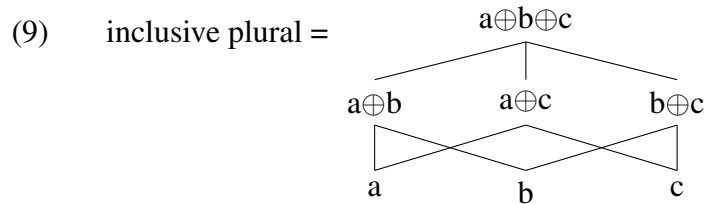
There is however a serious problem for this first hypothesis that plural in English is exclusive. Namely, the exclusive more-than-one interpretation is not present in negative sentences (and in other downward entailing contexts), as illustrated in (7). That is, the sentence in (7) typically does not convey the meaning that Malami did not buy more than one book but rather that he did not buy any book at all.

- (7) Malami didn't buy books.
 a. $\not\rightsquigarrow$ *Malami didn't buy more than one book*
 b. \rightsquigarrow *Malami didn't buy any book*

A solution that has been discussed in the previous literature is that the meaning of plural in English should include atomic individuals, i.e., it should be inclusive, as presented in (8) and (9)

below (see Yatsushiro et al. 2023 for experimental evidence):

- (8) a. $[[\text{book}]] = \{a,b,c\}$ SINGULAR
 b. $[[\text{book-s}]] = \{a, b, c, \{a,b\}, \{b,c\}, \{a,c\}, \{a,b,c\}\}$ INCLUSIVE PLURAL



The inclusive analysis of plural in English gives good results for negative sentences, predicting that (10) will convey the meaning in (10-a):

- (10) Malami didn't buy books.
 a. \rightsquigarrow *Malami didn't buy any book*

But then there is a question how the more-than-one meaning in upward-entailing contexts arises.

It follows that any analysis of the meaning of plural has to account for an observation that while plural count nouns convey the exclusive 'more-than-one' meaning in upward entailing contexts, they obtain the inclusive interpretation in downward entailing contexts (Sauerland 2003, Sauerland et al. 2005, Farkas and de Swart 2010, Grimm 2013, among others). Moreover, any analysis has to account for a dispreferred exclusive interpretation of plural under negation that sometimes arises, e.g., due to the prominence on the plural morphology, at least in English, as shown in (11):

- (11) Malami didn't buy books... he bought only one!

In order to account for these empirical observations, three main different analyses have been proposed in the previous literature: the ambiguity approach (e.g., Farkas and de Swart 2010, Martí 2020, Grimm 2013, see also Dawson 2024 in this volume for arguments for ambiguous plural in Tiwa), the implicature approach (e.g., Sauerland 2003, Spector 2007, Ivlieva 2013, Mayr 2015, see also Renans et al. 2018, 2020 and Tieu et al. 2020 for experimental evidence) and the homogeneity approach (Križ 2017). These approaches make however a crucial reference to the existence/availability of the inclusive plural. For example, while under the ambiguity approach, plural is ambiguous between the inclusive and exclusive meaning, under the implicature approach plural is semantically inclusive and the exclusive interpretation arises as a scalar implicature.

2.2 Languages with exclusive plural only?

A big issue in a cross-linguistic semantics of plural is the question of whether there is a language with exclusive plural only, i.e. a language in which plural nouns are invariably interpreted exclusively in upward and downward entailing contexts, as demonstrated in (12) and (13).

- (12) Malami bought books.

- (13) a. \rightsquigarrow *Malami bought more than one book* EXCLUSIVE PLURAL
 Malami didn't buy books.
 a. \rightsquigarrow *Malami didn't buy more than one book* EXCLUSIVE PLURAL
(he could have bought one book only)

It follows that the inclusive interpretation of plural should be completely unavailable in such a language, that is, (14-a) should not be a possible interpretation of (14) in exclusive-only languages:

- (14) Malami didn't buy books.
 a. \rightsquigarrow *Malami didn't buy any book* INCLUSIVE

There are some potential candidates for exclusive-only languages that have been discussed in the previous theoretical literature, e.g., Western Armenian (Bale and Khanjian 2014), Turkish (Bale and Khanjian 2014, Görgülü 2012), Skwxwú7mesh, Squamish (Bar-el 2008), Akan (Ahenkorah 2023), and Hausa (Součková 2011, Zimmermann 2021). Let me briefly comment on the type of evidence presented in support of the view that there is an exclusive-only plural in a language and/or a current state of the debate regarding the interpretation of plural in these languages.

Bale and Khanjian (2014) show in their paper that plural in Western Armenian is interpreted exclusively in positive sentences. They also provide the example in (15), which illustrates the interpretation of plural in the restrictor of quantifiers, another downward entailing context.

- (15) ?Amen mart vor bəzdig-**ner** uner vodk-i gajne-tsav.
 all person that child-PL had foot-DAT stand.up-PST
 'Everyone that had two or more children stood up.' EXCLUSIVE

Bale and Khanjian (2014) report that (15) is not really an acceptable sentence in Western Armenian but when forced to give an interpretation, the language consultants interpret the plural in the restrictor of the quantifiers exclusively. They also provide in the footnote an example of the negative sentence with plural count nouns, reported below in (16):

- (16) a. Bezdig-**ner** chi desah.
 child-INDF.PL not saw.1.SG
 'I didn't see children.'
 b. Voch bezdig-**ner** voch al gadu-**ner** desah.
 no child-INDF.PL no also cat-IND.PL saw.1.SG
 'I saw no children and no cats.' [Bale & Khanjian 2009, p. 16]

The data in (16) suggest that plural in negative sentences can obtain an inclusive interpretation. Bale and Khanjian (2014) claim however that the negation works in Western Armenian differently.

As for Turkish, it is one of the typically cited languages in the literature as having an exclusive-only plural. A recent analysis however puts forward that plural in Turkish is just like in English: it obtains the exclusive interpretation in upward entailing contexts and the inclusive interpretation in downward entailing contexts (Sağ 2017, 2019). This analysis is supported by the results of experimental studies (Renans et al. 2020).

Akan was claimed to have two types of plural: inclusive and exclusive-only (Ahenkorah 2023).

Owusu (2022) argues however that plural in Akan is not exclusive-only across the borders. Recent experimental studies are also in line with the exclusive/inclusive interpretation of plural, i.e., not exclusive-only (Bonney et al. 2023).

Summing up, it seems that more empirical work has to be done in order to find out whether there is a language with the exclusive-only plural. In the next section, I will carefully investigate the meaning of plural in Hausa, another language which has been claimed in the previous literature to have exclusive-only plural, both in the nominal and verbal domain.

3 Nominal plural in Hausa

Hausa is a Chadic language, spoken by ca. 35 million speakers, mostly in West Africa (it is spoken as first language in northern Nigeria and southern Niger but also in Ghana, Chad, and Sudan). If not marked otherwise, the data in the following sections come from author's direct elicitation with one Hausa native speaker in Berlin who speaks a Nigerian Hausa (April–October 2023) and a questionnaire elicitation via email with one speaker of Hausa in Ghana (April–June 2023). The methodology to elicit the data followed the guidelines of (Matthewson 2004). The tasks comprised mostly translations and acceptability judgments of sentences in the context.

3.1 Plural in upward entailing contexts

Hausa has a plethora of morphological plural markers: 40 surface forms, reduced to roughly 14 major classes (Součková 2011, Newman 2000, 2007, Jaggar 2001). Examples (17) and (18) exemplify the use of the morphologically plural and singular count nouns in upward entailing contexts.² In the context in which Halima planted only one flower, the singular count, but not the plural one, is acceptable, as shown in (17).

- (17) context: Halima wanted to plant flowers but her shovel broke down and she planted only one flower.
- a. #Halima tā shùkà hùrênnì.
Halima 3SG.F.COMPL plant flower.PL
'Halima planted flowers.'
- b. Halima tā shùkà hùrê.
Halima 3SG.F.COMPL plant flower.SG
'Halima planted a flower.'

By contrast, in the context in which Malami bought several books, as in (18), the plural count noun is preferred. (18-a) with a singular count noun is however still compatible with the context, as if when one buys more than one book then it entails that a singular book was bought.

²The following glosses are used in the paper: PRT = particle, DEF = definite determiner, COMPL = completive, 1,2,3 = 1st, 2nd, 3rd person, SG = singular, PL = plural, NEG = negation, F = feminine, M = masculine, MOD = modal, RED = reduplication. An example marked with '#' means that the example was judged to be unacceptable in the given context and we hypothesize that it is for semantic or pragmatic reasons; in the case of '?' the judgments were mixed. Finally, examples without any diacritics were judged as acceptable in the given context.

- (18) context: Malami went to the shop and bought several books.
- a. Malami yā sayi **littaafi**.
 Malami 3SG.M.COMPL buy book.SG
 ‘Malami bought a book.’
- b. Malami yā sayi **littaatafai**.
 Malami 3SG.M.COMPL buy book.PL
 ‘Malami bought books.’

Looking at the sentence with a plural count noun in (19), the language consultant offered a comment that (19) means that Malami read several books, he could not read just one book.

- (19) Malami yā kàrantà **littaatafai**.
 Malami 3SG.M.COMPL read book.PL
 ‘Malami read books.’
comment: he read more than one book, he couldn’t read just one book

These data suggest that plural in Hausa in upward entailing contexts gives rise to the more-than-one interpretation, i.e. plural is interpreted exclusively in these contexts. This finding is very much in line with previous descriptions and analyses of Hausa plural (Součková and Buba 2008, Součková 2011, Zimmermann 2008, 2021).

Before we jump into the interpretation of plural in downward entailing contexts, let me comment on the singular count nouns in Hausa. In the previous literature, it has been claimed that singular in Hausa is number-neutral, that is that its denotation comprises both atomic entities and all the pluralities formed out of them. The argument for this claim comes from the competition with plural, i.e., if plural is exclusive, then it is intuitively compelling to postulate a number-neutral singular (or vice versa: if singular is number-neutral, it is compelling to hypothesize an exclusive plural).³ The second line of arguments comes from the combination with numerals which I will discuss in details in Section 5.

The data in (20) challenge however the view that singular in Hausa is number-neutral. If it were number-neutral, then it should be possible to utter (20-a), contrary to fact. In order to refer to plurality of boys, a plural count noun must be used, as demonstrated in (20-b).

- (20) a. *Malami dà Yusuf **yaro** nè.
 Malami and Yusuf boy.SG be
 ‘Malami and Yusuf are boys.’
- b. Malami dà Yusuf **yara** nè.
 Malami and Yusuf boy.PL be
 ‘Malami and Yusuf are boys.’

To sum up, plural in Hausa gives rise to the more-than-one interpretation, i.e. it is interpreted exclusively in upward entailing contexts, and the view that singular is number-neutral is seriously challenged by the data presented in this section.

³See however (Dawson 2024) in this volume for an analysis of Tiwa: a language with a number-neutral singular and ambiguous plural. Crucially, Dawson shows that having a number-neutral singular does not entail exclusive plural.

3.2 Plural in downward entailing contexts

Let us now look at the interpretation of plural in downward entailing contexts: negation, antecedents of conditionals, and questions, starting with negation in (21):

- (21) context: Malami went to the shop and bought one book.
 a. #Malami bà yā sayi **littaatafai** ba.
 Malami NEG 3SG.M.COMPL buy book.PL NEG
 ‘Malami didn’t buy books.’
comment: not good, because he bought one book

Example (21-a) was not accepted in the context of (21) by the language consultant, suggesting that the inclusive interpretation of plural is available in Hausa (otherwise, the sentence in (21-a) would be perfectly acceptable in the context of (21)). Example (22) shows however that the exclusive interpretation of plural under negation can appear, just like in English:

- (22) context: Halima wanted to plant flowers but her shovel broke down and she planted only one flower.
 a. Halima bà tā shùkà **hùrènnì** ba.
 Halima NEG 3SG.F.COMPL plant flower.PL NEG
 ‘Halima didn’t plant flowers.’
comment: depending on the emphasis, it could either mean ‘Halima didn’t plant any flower at all’ or ‘Halima didn’t plant plural flowers, she planted only one’

An inclusive interpretation of plural is also detected in another downward entailing context, i.e. antecedents of conditionals, as demonstrated in (23):

- (23) context: Malami was supposed to feed pigs but he took too few food and he fed only one pig.
 a. Idan Malami yā ciyar dà **àladū**, sai mu gode masà.
 if Malami 3SG.M.COMPL feed with pig.PL, MOD we thank him
 ‘If Malami fed pigs, we should thank him.’
 Q: Shall we thank him?
 A: yes

The language consultants were presented with the context in (23) and the conditional sentence in (23-a). Subsequently they were asked whether we should thank Malami. The ‘yes’-answer given by the language consultants suggests that the inclusive interpretation of plural is available in the antecedent of conditionals.

Example (24) corroborates the view that the inclusive interpretation of plural is possible in downward entailing contexts in Hausa. Even though the answers were mixed (‘yes’ and ‘no’), the fact that ‘yes’ was also provided as an answer suggests that the inclusive interpretation of plural is available as well.

- (24) context: Malami was supposed to feed pigs but he took too few food and he fed only one

- (27) naa à'-'àikee sù
 1SG.PF RED-send them
 'I sent them to different places/to the same place at different times/to different places at different times.'
- (28) naa sàs-sàyi littàttfai
 1SG.PF RED-buy books
 'I bought many (different) books (on different occasions/in different bookstores.)'
 (from Součková and Buba 2008, p.134)

Pluractional markers can also convey other meanings, such as high individuation, large number of events, diversification or intensification (see e.g., Cusic 1981, Součková 2011). For illustration, intensification conveyed by reduplication is demonstrated in (29):

- (29) a. mun gáji
 1.PL.PF be.tired
 'We are tired.'
- b. mun gág-gàji
 1PL.PF RED-be.tired
 'We are all **very** tired.'
 (from Součková and Buba 2008, p.137)

I observe that pluractionals might also yield the multiplicity inference. For example, (25) gives rise to the inference that the table was kicked more than once. Even though the multiplicity inference in the nominal domain has been extensively investigated, it has not been researched in the verbal domain. This paper aims at filling in this gap and provide the first preliminary data on the interpretation of verbal plural in downward entailing contexts.

4.1 Verbal plural in upward entailing contexts

For sake of completeness, this section presents the data exemplifying the use of pluractionals in upward entailing contexts. The language consultant's judgments reveal that both (30-a) with reduplication and (30-b) without reduplication can be used in the situation in which the goat jumped many times. Example (30-a) however received a comment that it describes the context more accurately than (30-b).

- (30) context: Our goat is really active! He jumped over the fence many times last week.
- a. Bùnsurunmù ya tsa-tsallàkè dangaa.
 goat.our 3SG.M.COMPL RED-jump fence
- b. Bùnsurunmù ya tsallàkè dangaa.
 goat.our 3SG.M.COMPL jump fence
 'Our goat jumped over the fence.'
comment: all good but (a) is better than (b)

By contrast, in the context in which the goat jumped only once, (31-b) is still acceptable but (31-a), with reduplication, not.

- (31) context: Our goat became really calm these days. He jumped over the fence only once last week.
- a. #Bùnsurunmù ya **tsa**-tsallàkè dangaa.
goat.our 3SG.M.COMPL RED-jump fence
- b. Bùnsurunmù yaa tsallàkè dangaa.
goat.our 3SG.M.COMPL jump fence
'Our goat jumped over the fence.'

These data suggest that pluractionals in Hausa give rise to more-than-one meaning in upward entailing contexts.

4.2 Verbal plural in downward entailing contexts

Let us now consider the interpretation of pluractionals in downward entailing contexts, starting from *if*-clauses. The language consultants were presented with the context in (32) together with a conditional sentence in (32-a), containing a pluractional in its antecedent. The task of the language consultants was to decide whether one should call us or not.

- (32) context: Malami kicked the table only once.
- a. Idan Malami yā **shùs**-shùuri teebùr, sai ka kirā mù.
If Malami 3SG.M.COMPL RED-kick table MOD you call us
'If Malami kicked the table, you should call us.'
- Q: Should you call us?
A: yes

A positive answer given by the language consultants suggests that atomic/singular events are also present in the denotation of the plural verb. This aligns with the judgments given for the next test. The language consultants were asked whether Malami kicked the table (using the pluractional reduplicated form) in the context in which he kicked a table only once, as illustrated in (33).

- (33) context: Malami kicked the table only once.
Malami yā **shùs**-shùuri teebùr?
Malami 3SG.M.COMPL RED-kick table
'Did Malami kick the table?'
- A: yes

There were however mixed answers regarding the interpretation of plural verbs under negation: the sentence in (34-a) was judged to be acceptable in the context of (34), which specifies that Malami did kick the table, and the sentence in (35-a) was judged to be acceptable in the context of (35), which makes it clear that Malami did not feed the pigs at all. These data might suggest that the dispreferred exclusive reading under negation might also arise in the verbal domain, as suggested by the judgments in (34). The inclusive interpretation of plural verb under negation is however also available, as suggested by the data in (35).

- (34) context: Malami kicked the table only once.

- a. Malami bai **shùs**-shùuri teebùr ba.
 Malami NEG RED-kick table NEG
 ‘Malami didn’t kick the table.’
comment: ok if he kicked the table only once

(35) context: Malami was supposed to feed pigs every day last month but he did not feed them even once! His brother took over.

- a. Malami bai **ci**-ciyar dà àladū ba.
 Malami NEG RED-feed with pig.PL NEG
 ‘Malami didn’t feed pigs.’

Summing up, verbal plural gives rise to the more-than-one/exclusive meaning in upward entailing contexts. Moreover, the first preliminary data seem to suggest that an inclusive interpretation of verbal plural is available in downward entailing contexts.⁶ While definitely more empirical work is needed on the interpretation of nominal and verbal plural in upward and downward entailing contexts, the data presented in this paper challenge the existing view that plural in Hausa is exclusive-only and asks for a careful investigation of this issue.

5 Number morphology and numerals

This section discusses the interaction of plural morphology with numerals. It is an important issue because the possibility of combining the numerals with morphologically singular count nouns has been given in the previous literature as an argument in favor of the exclusive plural in Hausa (and inclusive singular). Since I argue in this paper that up to now we have no evidence in favor of the exclusive-only denotation of plural in Hausa, let me comment on this issue.

5.1 Number morphology and numerals across languages

With respect to the combination with numerals, one can distinguish three types of languages with overt plural morphology: (i) languages such as English, in which nouns obligatory take plural forms when combined with numerals bigger than one:

- (36) a. two book*(s) ENGLISH
 b. one book(*s)

(ii) languages such as Turkish and Hungarian, in which nouns obligatory take their bare form when combined with numerals bigger than one:

- (37) TURKISH:
 iki kitap-(*lar)
 two book-PL
 ‘two books’ (Renans et al. 2020, p.339)

⁶In the paper, I showed the data with the iterative reading of the sentences with pluractionals. My preliminary data on other readings suggest that other readings also show a discussed contrast in the interpretation in upward and downward entailing contexts.

- (38) HUNGARIAN:
három guerek-(*ek)
three child-PL
'three children' (Farkas and de Swart 2010, p.10)

(iii) and languages such as Western Armenian and Halkomelem, in which nouns take both bare and plural forms when combined with numerals bigger than one:⁷

- (39) WESTERN ARMENIAN:
yergu dōgha-(ner)
two boy-PL
'two boys' (Bale et al. 2010, p.593)

- (40) HALKOMELEM:
a. te lh̀w swòweles
DET three boy.PL
b. te lh̀w swiweles
DET three boy
'the three boys' (Wiltschko 2008, p. 642)

Looking at Hausa, both singular and plural forms are allowed with numerals bigger than one, as demonstrated in (41):⁸

- (41) HAUSA:
a. **littaafi** biyu
book.SG two
b. **littaatafai** biyu
book.PL two
'two books'

However, Hausa differs in interesting ways from Halkomelem, and Western Armenian (and Tiwa, see Dawson 2024). First, while in Tiwa and Halkomelem, plural form is fully optional, it seems that in Hausa plural form is pretty obligatory without numerals and quantifiers (when referring to plural entities). Second, while in Tiwa, Halkomelem and Western Armenian, singular form is a general number, it can refer both to singular and plural entities, the tests given as argument for this claim in Tiwa and Western Armenian gave different results in Hausa, see data in (20). Thus there is no conclusive evidence for a general number/number-neutrality in Hausa, see section 3.1 for discussion and data.

⁷see also (Dawson 2024) in this volume on the semantic of plural in Tiwa, a language in which numerals can also combine with both singular and plural form of a noun.

⁸There is an optional classifier *guda* in Hausa. I am leaving its analysis for future research.

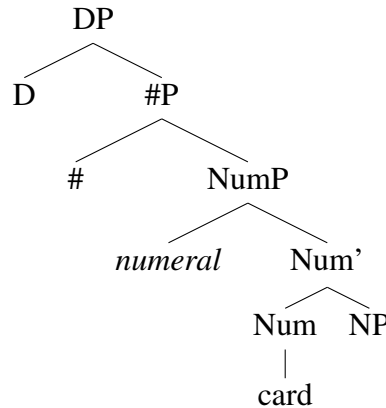
5.2 Towards an analysis

In the previous literature, several analyses have been proposed to account for the interaction between plural/singular morphology and numerals. Ionin and Matushansky (2004, 2006, 2018) argue that numerals combine with singulars and the plural marker *-s* in English is an agreement (see also Alexiadou 2019, Krifka 1989, Sağ 2017, 2019). It is however difficult to see how this analysis could be extended to Hausa, since in Hausa both morphologically singular and plural count nouns can combine with numerals.

Bale et al. (2010) developed an analysis that account for the possibility of combining the numerals with both morphologically singular and plural count nouns. However, they make crucial reference in their analysis to the number-neutrality of singular nouns, which is not found in Hausa and thus an extension of this analysis to the Hausa language is at least not straightforward.

The most promising in light of Hausa data seems an analysis of number morphology and numerals developed by Scontras (2014, 2022), see also Martí (2020). Scontras argue for a following structure of the DP:

(42)



Num determines the function of the numerals (cardinal, ordinal); card is employed for counting. The number morphology is assigned in the # (after numerals combine with nouns). The role of singular morphology is to check for atomicity. If every element in the denotation of NumP is atomic, then the singular morphology is assigned, if not, then the plural morphology is assigned. Crucially, the checking for atomicity is a presupposition and the choice of the singular/plural morphology is regulated by Maximize Presupposition (Heim 1991). The semantics of SG and PL is given in (43):

(43) #'s semantics:

- a. $[[SG]] = \lambda P : \forall x \in P[\mu(x) = 1].P$
- b. $[[PL]] = \lambda P.P$

(Scontras 2022, (13))

This accounts for the English pattern, in which numerals bigger than one require plural morphology. Namely, since there are only atomic elements in the denotation of *one NP*, a singular morphology is assigned with the numeral *one*. By contrast, there are no atomic elements in the denotation of *two NP* and therefore plural morphology is assigned with numerals bigger than one.

To account for the Turkish pattern, in which numerals require singular, Scontras (2022) argues

for the same structure of the DP and the same semantics of #. The difference is located in the definition of atomicity. While the English SG checks for pure atoms, the Turkish SG checks for impure P-atoms, defined as in (44-c).

- (44) a. $\llbracket \text{SG} \rrbracket = \lambda P : \forall x \in P [\mu_{P\text{-atom}}(x) = 1].P$
 b. $\llbracket \text{PL} \rrbracket = \lambda P.P$
 c. $\mu_{P\text{-atom}}(y)$ is defined only if $y \in P$; when defined
 $\mu_{P\text{-atom}}(y) = |\{x \in P : x \leq y \wedge \neg \exists z \in P [z < x]\}|$
 ‘every member of the predicate has no proper parts that are themselves member of this predicate; every member of this predicate measures 1 P-atom’
 (Scontras 2022, (45)-(46))

Under this definition of atomicity, also the elements in the denotation of nouns combined with numerals bigger than one are (impure) atoms (e.g., all the elements in the denotation of *two dogs* are impure P-atoms). Therefore, singular morphology is required with numerals bigger than one. In order to account for Western Armenian pattern, Scontras (2022) proposes that in this language both strategies are available and therefore numerals can combine with both morphologically singular and plural nouns.

Looking at Hausa data, one could adopt the view that in Hausa, just as in Western Armenian, both strategies are available, i.e., that SG can make reference to both pure and impure P-atoms. It means however that SG would be ambiguous between referring to pure and impure P-atoms which is not really attractive from the theoretical perspective.

I would like to entertain another possibility, namely that actually SG in Hausa is not ambiguous: it always refers to P-atoms and thus in principle only the Turkish pattern is available in Hausa. The difference is however that P-atomicity is not a part of presupposition but an entailment. If this is so, then it follows that the Maximize Presupposition does not regulate the choice of SG vs. PL. Therefore, the optionality in choice of SG vs. PL with numerals is predicted but there is no optionality predicted with respect to bare nouns. If this set up is on a right track, it would suggest that languages differ with respect to whether SG makes reference to pure or impure P-atoms and whether SG and PL competes.

6 Summary

A take-home message is that the question of whether there are languages with exclusive-only plural is far from being set up. The novel data from Hausa coming from the behavior of plural in downward entailing contexts seriously challenge the view that has been made in the previous literature that Hausa is a language with an exclusive-only plural.

Especially interesting in the debate on the exclusive vs. inclusive plural is the meaning of pluractionals, the verbal plural. Even though all existing theories of pluractionality actually predict the existence of an exclusive-only plural in the verbal domain, the first preliminary data on the behavior of pluractionals in downward entailing contexts in Hausa suggest that inclusive interpretation of verbal plural is available. This in turn asks for a revision of the analysis of pluractionality.

Saying that, more cross-linguistic empirical work is needed in the domain of plurality across

domains. In this respect, especially interesting are Salish languages which mark both the nominal and verbal plural by reduplication and which have been argued to make the same contribution in both domains (see e.g. Wiltschko 2008). Therefore, it would be really interesting to examine their behavior in downward entailing contexts and to see whether they really behave alike.

References

- Ahenkorah, C. (2023). The syntax-semantic analysis of Akan plural morphology.
- Alexiadou, A. (2019). Morphological and semantic markedness revisited: The realization of plurality across languages. *Zeitschrift für Sprachwissenschaft*, 38(1):123–154.
- Bale, A., Gagnon, M., and Khanjian, H. (2010). Cross-linguistic representations of numerals and number marking. In Li, N. and Lutz, D., editors, *Semantics and Linguistic Theory (SALT) 20*, pages 582–598.
- Bale, A. and Khanjian, H. (2014). Syntactic complexity and competition: The singular-plural distinction in Western Armenian. *Linguistic Inquiry*, 45(1):1–26.
- Bar-el, L. (2008). Verbal number and aspect in Skwxwú7mesh. *Recherches linguistiques de Vincennes*, (37):31–54.
- Bonney, R. D., Ladore, M., Struck, A., Tanihu, J., Laryea, H., and Abdulai, K. (2023). Exploring plural making in Akan: An experimental analysis.
- Chierchia, G. (1998). Reference to kinds across languages. *Natural Language Semantics*, 6(4):339–405.
- Cusic, D. D. (1981). *Verbal Plurality and Aspect*. PhD thesis, Stanford University.
- Dawson, V. (2024). Inclusive plural in a ‘general number’ language. In *Proceedings of TripleA 10*, pages –.
- Farkas, D. F. and de Swart, H. (2010). The semantics and pragmatics of plurals. *Semantics and pragmatics*, 3:1–54.
- Görgülü, E. (2012). *Semantics of nouns and the specification of number in Turkish*. PhD thesis, Simon Fraser University.
- Grimm, S. (2013). Plurality is distinct from number-neutrality. In *Proceedings of the 41st Meeting of the North East Linguistic Society (NELS) 41*, pages 247–258.
- Harbour, D. (2014). Paucity, abundance, and the theory of number. *Language*, pages 185–229.
- Heim, I. (1991). Artikel und definitheit. *Semantik: ein internationales Handbuch der zeitgenössischen Forschung*, 6:487–535.
- Ionin, T. and Matushansky, O. (2004). A singular plural. In *Proceedings of WCCFL*, volume 23, pages 399–412.
- Ionin, T. and Matushansky, O. (2006). The composition of complex cardinals. *Journal of semantics*, 23(4):315–360.
- Ionin, T. and Matushansky, O. (2018). *Cardinals: The syntax and semantics of cardinal-containing expressions*, volume 79. MIT Press.
- Ivlieva, N. (2013). *Scalar implicatures and the grammar of plurality and disjunction*. PhD thesis, Massachusetts Institute of Technology.
- Jaggard, P. J. (2001). *Hausa*. London Oriental and African Language Library 7. Benjamins, Amsterdam.

- Krifka, M. (1989). Nominal reference, temporal constitution and quantification in event semantics. volume 75, pages 9783110877335–005.
- Križ, M. (2017). Bare plurals, multiplicity, and homogeneity. Ms ENS.
- Lasersohn, P. (1995). *Plurality, Conjunction and Events*. Kuwer Academic Publishers.
- Link, G. (1983). The logical analysis of plurals and mass terms: A lattice-theoretical approach. In Baeuerle, R., Schwarze, C., and Stechow, A. v., editors, *Meaning, Use and Interpretation of Language*. DeGruyter.
- Martí, L. (2020). Numerals and the theory of number. *Semantics & Pragmatics*, 13(3):1–57.
- Matthewson, L. (2004). On the methodology of semantic fieldwork. *International journal of American linguistics*, 70(4):369–415.
- Mayr, C. (2015). Plural definite NPs presuppose multiplicity via embedded exhaustification. In *Semantics and Linguistic Theory (SALT) 25*, pages 204–224.
- Newman, P. (1990). *Nominal and verbal plurality in Chadic*. Dordrecht: Foris Publications.
- Newman, P. (2000). *The Hausa Language: An Encyclopedic Reference Grammar*. Yale University Press, New Haven.
- Newman, P. (2007). *Hausa*. Eisenbrauns.
- Newman, P. (2019). Nominal and verbal plurality in Chadic. In *Nominal and Verbal Plurality in Chadic*. De Gruyter Mouton.
- Owusu, A. P. (2022). *Cross-Categorical Definiteness/Familiarity*. PhD thesis, Rutgers University.
- Pawlak, N. (1975). The semantic problems of ‘intensive’ forms in Hausa. *African Bulletin*, (23):139–149.
- Renans, A., Romoli, J., Makri, M.-M., Tieu, L., de Vries, H., Folli, R., and Tsoulas, G. (2018). The abundance inference of pluralised mass nouns is an implicature: Evidence from Greek. *Glossa: A Journal of General Linguistics*, 3(1):1–34.
- Renans, A., Sağ, Y., Ketz, N., Tieu, L., Folli, R., Tsoulas, G., de Vries, H., and Romoli, J. (2020). Plurality and cross-linguistic variation: An experimental investigation of the Turkish plural. *Natural Language Semantics*, 28(4):307–342.
- Sauerland, U. (2003). A new semantics for number. In *Semantics and linguistic theory*, volume 13, pages 258–275.
- Sauerland, U., Andersen, J., and Yatsushiro, K. (2005). The plural is semantically unmarked. In Kepser, S. and Reis, M., editors, *Linguistic Evidence: Empirical, Theoretical, and Computational Perspectives*, pages 413–434. Mouton de Gruyter.
- Sağ, Y. (2017). The semantics of numeral constructions in Turkish. In *Sinn und Bedeutung (SuB) 22*, pages 307–324.
- Sağ, Y. (2019). *The semantics of number marking: reference to kinds, counting, and optional classifiers*. PhD thesis, Rutgers University.
- Scontras, G. (2022). On the semantics of number morphology. *Linguistics and Philosophy*, 45(5):1165–1196.
- Scontras, G. C. (2014). *The semantics of measurement*. PhD thesis.
- Součková, K. (2011). *Pluractionality in Hausa*. PhD thesis, LOT.
- Součková, K. and Buba, M. (2008). Intensive plurality. *Linguistics in the Netherlands*, pages 133–144.
- Spector, B. (2007). Aspects of the pragmatics of plural morphology: On higher-order implicatures.

- In Sauerland, U. and Stateva, P., editors, *Presupposition and Implicature in Compositional Semantics*. Palgrave.
- Tieu, L., Bill, C., Romoli, J., and Crain, S. (2020). Testing theories of plural meanings. *Cognition*, 205:104307.
- Tovena, L. M. and Kihm, A. (2008). Event internal pluractional verbs in some Romance languages. In Tovena, L. M., editor, *Aspect et pluralité d'événements*, volume 37 of *Recherches Linguistiques de Vincennes*, pages 31–54.
- van Geenhoven, V. (2004). For-adverbials, frequentative aspect, and pluractionality. *Natural Language Semantics*, 12:135–190.
- van Geenhoven, V. (2005). Atelicity, pluractionality, and adverbial quantification. In Verkuyl, H. J., de Swart, H., and van Hout, A., editors, *Perspectives on Aspect*, pages 107–124. Springer.
- Wiltschko, M. (2008). The syntax of non-inflectional plural marking. *Natural language & linguistic theory*, 26(3):639–694.
- Yatsushiro, K., Geçkin, V., Harmati-Pap, V., Alexiadou, A., and Sauerland, U. (2023). The plural is unmarked: Evidence from turkish, hungarian and german. *Glossa: a journal of general linguistics*, 8(1).
- Zimmermann, M. (2003). Pluractionality and complex quantifier formation. *Natural Language Semantics*, 11:249–287.
- Zimmermann, M. (2008). Quantification in Hausa. *Quantification: Universals and variation*, 64:415–475.
- Zimmermann, M. (2021). Verbal number in Chadic, with special reference to Hausa. In Hofherr, P. C. and Doetjes, J., editors, *The Oxford Handbook of Grammatical Number*, pages 597–625.