Question word distributivity[1]

Mary MORONEY — *University of Rochester*

Abstract. The Shan (Kra-Dai) question word for 'which' *lǎj* (or *lǎɯ* in some varieties) can also be used to have a distributive interpretation similar to *each*. This interpretation exists in addition to the indeterminate pronoun meanings, e.g., as a negative polarity item, question word, or universally quantified expression, such as are found in Japanese. This paper describes the distributive construction and indeterminate pronouns in Shan and develops a unified account for *lǎj*. The same Hamblin semantics that can account for indeterminate pronouns is extended to the distributive case with the incorporation of Neo-Davidsonian event semantics. This identifies a previously unseen connection between indeterminate pronouns and distributivity.

Keywords: distributivity, indeterminate pronouns, multiple-*wh* questions, Shan, Kra-Dai.

1. Introduction

Distributivity is a means of grammatically indicating what amount of some object or portion is associated with a situation or event involving another object or portion. It involves three key components, (i) a plural KEY representing one plurality of individuals/portions, (ii) a SHARE predicate, and (iii) overt or covert distributive morphology that connects the SHARE to a distribution of the KEY (terminology from Champollion, 2019).

Languages cross-linguistically employ several morphological strategies for marking distributivity. These distributive morphemes can include determiners associated with the distributive key, as in (1a), adverbs that modify the share predicate, as in (1b), and morphemes that attach to the quantity in the distributive share, as in (1c). English *each* can do any of these, though each option is associated with different syntactic restrictions. In some languages, markers of distributivity appear on the distributive share quantity. For example, Kaqchikel uses reduplication of the numeral expression in the share, as in (2), and Bengali uses a morpheme associated with the numeral in the share, as in (3).

(1) a. **Each** of the children built a sand castle. DETERMINER (Champollion, 2019: (3))
    b. The children **each** built a sand castle. ADVERB (Champollion, 2019: (4a))
    c. The children built one sand castle **each**. ADNOMINAL (Champollion, 2019: (4b))

(2) Ri ak'wal-a' ni-Ø-ki-tij **ox-ox** wäy.
    The child-PL ICP-A3S-E3P-eat three-RED tortilla
    'The children are eating three tortillas each.' (Kaqchikel, Henderson, 2012: (22b))

(3) Robi ar Rina du-ṭo-**kore**-bakšo boeche.
    Robi and Rina two-CL.-do.PFV-box carried
    'Robi and Rina carried different sets of two boxes.' (Bengali, Guha, 2021: (3))

[1] Special thanks to ၸၢႆးသႅင်ယွတ်ႈ, ၸၢႆးႁၢၼ်သႅင်, ၼၢင်းမူလ်ၵူး and Hsai Mork Hom who provided the Shan data. Thanks also to members of the linguistics department at the University of Rochester for their feedback. ၶွပ်ႈၸႂ်ၼမ်ၼမ်ၶႃႈ!

1.1. Distributive use of *lăj*

Shan, a Kra-Dai language of Myanmar, expresses distributivity primarily using a structure that is unique for three reasons (i) the distributive morpheme has the form, *lăj*, which is homophonous with the question word meaning 'which', (ii) the distributive morpheme associates with the amount connected to the KEY, and (iii) since the distributive morpheme is marked on the distributive key, it is possible to distribute to non-atomic portions of the key.

In order to indicate a distributive reading, a form such as in (4) can be used.[2] The numeral and classifier, *nɯŋ kɔ̂* 'one person', that precede *lăj* indicate each distribution of the KEY, *lukhén sǎam kɔ̂ nâj* 'these three students', that is associated with the following SHARE predicate, which also must include a numeral and classifier.[3]

(4) lukhén sǎam kɔ̂ nâj, nɯŋ kɔ̂ **lăj** sɯ̂ pâplik sɔ̌ŋ pâp lɛ
student three CLF.HUM this one CLF.HUM LAJ buy book two CLF.BOOK and
kǎmtsɯ́m haa lɛm.
pen five CLF.THIN
'These three students bought two books and five pens per person.'
(Context: There are 3 students, 6 books, and 15 pens total.)

1.2. Indeterminate pronoun use of *lăj*

The word *lăj* also appears as a question word with a use similar to 'which' as in (5). This form also obligatorily appears with a classifier *ti* 'place' but without a numeral. Also, no numeral must follow *lăj* in contrast to the distributive construction.

(5) káa nân kút ti **lăj**?
car that park place LAJ
'**Where** (literally, **which** place) is that car parked?'

This morpheme is not just a question word. It is an indeterminate pronoun, a term first used by Kuroda (1965) to describe Japanese. In Shan, indeterminate pronouns can have for example a universal meaning, (6), or negative polarity item use, (7), in addition to the question use.

(6) mɤ **lăj** kɔ…
when LAJ PRT
'whenever …'

(7) ʔàm …táaŋ **lăj**
NEG way LAJ
'not …anywhere'

Cross-linguistically, we see distributivity (i) marked on the share, (ii) associated with the distributive share amount, or (iii) marked on the key with an atomic interpretation, like *each child*. This paper analyzes the compositional semantics of distributivity in a language like Shan that marks distributivity on an amount specified for the key. Additionally, the fact that this distributive morpheme has the same form as an indeterminate pronoun prompts a look at the connec-

[2]The Shan data presented here come from four speakers of Southern Shan.
Glossing conventions: 1: first person, 2: second person, ANML: animal, CLF: classifier, COMP.REL: relative clause complementizer, FEM: feminine, G: general, HUM: human, MASC: masucline, NEG: negation, NMLZ: nominalizer, PFV: perfective, PRT: particle, PL: plural, Q: question, RND: round.

[3]Shan has NUMERAL CLASSIFIERS in the sense of Aikhenvald (2000), which are obligatory with numerals and most quantifiers. Throughout this paper they will be referred to as simply 'classifiers'.

tion between these two constructions. Here, I argue that we can assume the same underlying semantics for *lăj* in the distributive and indeterminate pronoun uses.

The structures of the indeterminate pronoun and distributive uses of *lăj* are presented in Section 2. The distributive semantics that must be accounted for is discussed in Section 3. Section 4 introduces a Hamblin semantics for indeterminate pronouns, and Section 5 extends this account of indeterminate pronouns to account for distributive uses. Section 6 concludes.

2. Structure of distributive and indeterminate pronoun constructions

Shan is a Kra-Dai language spoken in Myanmar by approximately 4.7 million speakers (Ethnologue, 2021). It is an SVO language, and as (8) shows, the word order in nominal expressions is N Adj Num Clf Clf.Pl Dem.[4]

(8) háw khaj sûɪ măa tŏ-jàj săam tŏ tsɤ́ nân.
1 want buy dog body-large three CLF.ANML CLF.PL that
'I want to buy that group of three big dogs.'

(9) and (10) demonstrate the difference between a distributive use of *lăj* and a multiple-*wh* use of *lăj*.[5] For a distributive example such as (9), the plural KEY, *săam tsúm nâj* 'these three groups', is followed by a numeral and classifier, *nɯŋ tsúm* 'one group', that indicates the distributive amount for the key. This is then followed by the morpheme *lăj*, which precedes the SHARE, which contains another numeral-classifier modified nominal, *tŏnâp sŏŋ ʔăn* 'two numbers'.

(9) săam tsúm nâj, **nɯŋ tsúm lăj** lɤk tŏnâp sŏŋ ʔăn.
three group this one group LAJ choose number two CLF.G
'Those three groups, each group picked two numbers.'
(Context: 3 groups of students play a game. Each group picks two numbers (1–9).)

In a multiple-*wh* question involving *lăj*, such as (10), the two uses of *lăj* 'which' are part of the subject and direct object. This *wh*-expression must follow a classifier, which is obligatory with most numerals and quantifiers. In distinction with the distributive use, the question word *lăj* does not appear with a numeral. This makes the two uses easy to distinguish.

(10) **tsúm lăj** lɤk tŏnâp **ʔăn lăj**?
group LAJ choose number CLF.G LAJ
'Which group chose which number(s)?'
(Context: 3 groups of students play a game. Each group picks two numbers (1–9).)

2.1. Indeterminate pronoun structure

The distributive and indeterminate pronoun uses of *lăj* are clearly distinguishable by form. The indeterminate construction must appear with a classifier and without the numeral 'one' or any numeral, as in (11). This is similar to the Shan equivalent of 'sole/lone/only one', as in (12), which also does not have the numeral 'one'. When describing a singular individual, demonstratives often appear with a classifier and no numeral, as in (13), but the classifier can

[4]The plural classifier will not be discussed further here, but see Moroney (2021: §3.4.4) for more details.

[5]For nouns that describe a group or container, such as *tsúm* 'group', the noun that normally precedes the numeral and classifier is often elided.

be omitted, leaving plurality unspecified, as in (14), or higher numerals can be included, as in (15).

(11) kón (#nɯŋ) kɔ̂ lǎj
person one CLF.HUM LAJ
'which person'

(12) kón kɔ̂ lěw
person CLF.HUM alone
'sole person'

(13) kón (#nɯŋ) kɔ̂ nâj
person one CLF.HUM this
'this person'

(14) kón nâj
person this
'this person/these people'

(15) kón sɔ̌ŋ kɔ̂ nâj
person two CLF.HUM this
'these two people'

2.2. Distributive structure

There are three structures that can be used to indicate distributivity in Shan, similar to the three found in English. Two of these appear before the clause predicate and are not easily distinguished. As shown in (4), repeated below, the distributive morpheme *lǎj* can appear in an adverbial position between an overt plural subject or topic, *lukhén sǎam kɔ̂ nâj* 'these three students', and a predicate that includes a quantity, such as in the direct object *pâplik sɔ̌ŋ pâp lɛ kǎmtsʊ́m haa lɛm* 'two books and five pens'. Alternatively, the distributive expression could form a constituent with the subject as in (16). Here, there is already a plurality of teachers that have been described in the preceding context. This might be treated as the determiner distributive, equivalent to *each teacher*. This differs from (4) in that the understood subject is not a separate phrase from the distributive expression, and an overt noun forms a constituent with the distributive expression.

(4) lukhén sǎam kɔ̂ nâj, nɯŋ kɔ̂ **lǎj** sʊ̂ɪ pâplik sɔ̌ŋ pâp lɛ
student three CLF.HUM this one CLF.HUM LAJ buy book two CLF.BOOK and
kǎmtsʊ́m haa lɛm.
pen five CLF.THIN
'These three students bought two books and five pens per person.' ADVERBIAL
(Context: There are 3 students, 6 books, and 15 pens total.)

(16) khúsɔ̌n nɯŋ kɔ̂ **lǎɯ** tsɔjthɛ̌m lukhén sǎam kɔ̂.
teacher one CLF.HUM LAJ help student three CLF.HUM
'Each teacher helped three students.' DETERMINER

The third position for the distributive is to the right of the clause, immediately before the quantity that describes the distributive share, as in (17). This is similar to the adnominal structure in English. (18) has the same determiner structure as (16) but the same meaning as (17). The position of the distributive expression do not seem to affect the interpretation, but some speakers prefer the version where it appears later in the clause, i.e., (17).

(17) lukhén laj pâplik **sǎam kɔ̂ lǎj** sɔ̌ŋ pâp.
student get book three CLF.HUM LAJ two CLF.BOOK
'Each group of three students gets two books.' ADNOMINAL

(18) lukhén **săam kô** **lăj** laj pâplik sɔ̆ŋ pâp.
student three CLF.HUM LAJ get book two CLF.BOOK
'Each group of three students gets two books.' DETERMINER

In Shan, numeral-classifier expressions can also appear in a floated position that follows clause final aspect markers, and the distributive construction can appear in the same position.[6] The floated distributive construction in (19), which follows the aspect marker *jâw*, consists of a modified numeral-classifier expression where the distributive expression, *nɯŋ kô lăj* 'each one/per person', modifies the floated direct object numeral-classifier expression. This fits with the account of adnominal expressions from Zimmermann (2002) where the phrase containing the distributive morpheme in German adjoins to the distributive share.

(19) lukhén táŋseŋ laj màakmoŋ jâw nɯŋ kô lăj sɔ̆ŋ hòj.
student all get mango PFV one CLF.HUM LAJ two CLF.RND
'All the students have gotten mangoes, two per person.'
(Context: There are 9 students and 18 mangoes.)

This adnominal account fits well with (20), where *nɯŋ kô lăj sɔ̆ŋ ʔăn* 'two per person', intervenes between the noun *kăanwaan* 'homework' and a relative clause modifying that noun. This would make sense if the distributive *nɯŋ kô lăj* 'per person' modifies *sɔ̆ŋ ʔăn* 'two'.

(20) khúsɔ̆n săam kô nân tsɤ́-khăw ʔɔ̆nkăn thátsaang kăanwaan nɯŋ kô
teacher three CLF.HUM that PL-3PL together examine homework one CLF.HUM
lăj sɔ̆ŋ ʔăn tsɤ́ ʔăn lukhén khăw sòŋ máa nân.
LAJ two CLF.G CLF.PL COMP.REL student 3PL send come that
'Those three teachers, examined the homework, two each, that their students sent in.'

The distributive construction can be used to distribute over times and places not otherwise expressed in the sentence, as in (21). Here, the distributive expression is clause-final. The portion associated with the key is *nɯŋ lɤ̆n* 'one month', and the total number of months is not specified. This is followed by the distributive morpheme *lăj*, and the distributive morpheme is followed by the amount associated with the share *nɯŋ pɔk* 'one time', an adverbial expression.

(21) háw kwàa thóp mɔ̆jăa nân nɯŋ lɤ̆n **lăj** nɯŋ pɔk.
1 go meet doctor that one month LAJ one time
'I go see that doctor once each month.'

The distributive construction and indeterminate-*which* both make use of CLASSIFIER+*lăj*. Their structural differences are summarized in (22) and (23). The difference in their structures can be attributed to a difference in selection: the distributive *lăj* selects for a numeral-classifier expression while the indeterminate pronoun *lăj* selects for just the classifier expression. Additionally, I propose that the indeterminate pronoun functions much like a determiner, while the distributive *lăj* must be licensed by an overt distributive operator.

(22) **Distributive** $[(N_1)\ Num_1\ Clf\ lăj]_{KEY\text{-}AMOUNT} \ldots [(N_2)\ Num_2\ Clf]_{SHARE\text{-}AMOUNT}$
'Num_2 N_2 per Num_1 N_1'

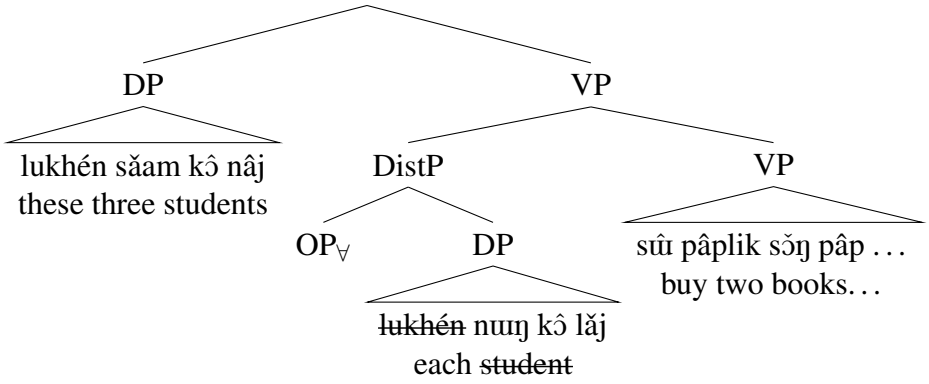
(23) **Indeterminate pronoun** [(N) Clf *lăj*]
'Which N'

[6] Chaiphet (2023) analyzes a similar pattern in Thai as subextracted quantifiers moving to a focus position.

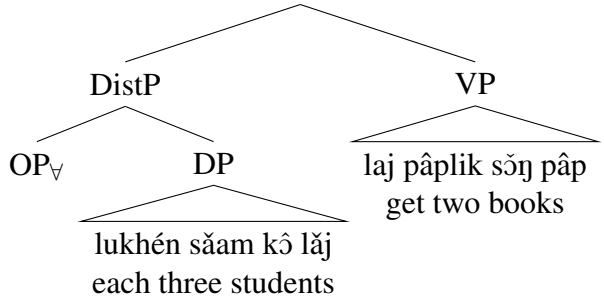
I will assume the constituent containing *lăj* is always a DP and propose that the noun is elided when it is not visible because it is anaphoric to an overt plural KEY. This is similar to adverbial *each* in English, which includes an overt noun only when the distributive key is not an argument of the verb, e.g., *We won **each time*** vs *The students **each** won*. However, this is not exactly the same as the English case, since clear anaphoric determiner uses of *each* in English often include the pronominal *one*, e.g., ***Each** (**one**) won*, which is not allowed in the adverbial use.

The structures for the ADVERBIAL distributive in (4) and the DETERMINER distributive in (18) are proposed to be (24) and (25), respectively. In both cases, the morpheme *lăj* in the context of a numeral-classifier indicates that a distributive operator is present in the structure, and *lăj* has the same syntactic position in all distributive cases. The description of these as being adverbial, determiner, or adnominal expressions describes the structural position of the NUM-CLF *lăj* distributive phrase. The anaphoric subject or topic is only overtly expressed in (24), while in (25) the distributive expression has some salient plural antecedent, possibly from a previous sentence. For the adverbial construction, the distributive phrase modifies a verb phrase, and in the determiner construction, it functions as an argument of the verb phrase.

(24) ADVERBIAL DISTRIBUTIVE

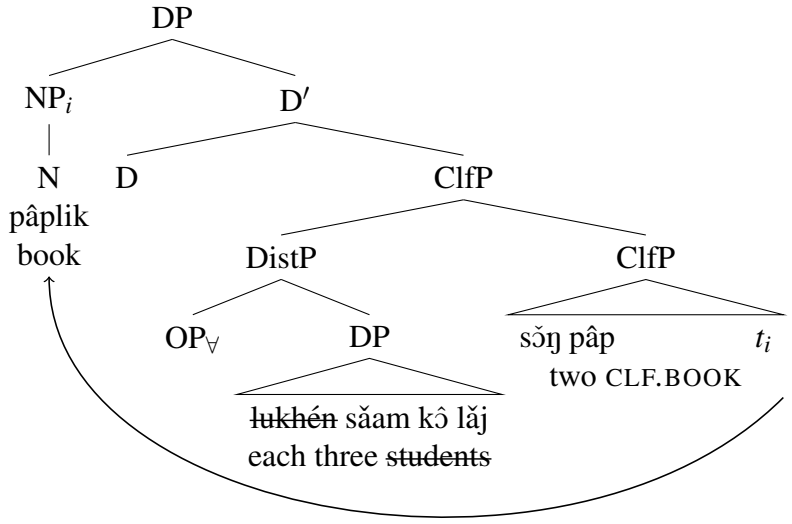


(25) DETERMINER DISTRIBUTIVE



Shan nominal syntax is similar to Thai. Both are Southwestern Tai languages. The internal structure of Shan nominals will not be presented here due to space limitations, but see Simpson (2005); Jenks (2011); Chaiphet (2023); and others for details on Thai. Most accounts of Thai propose movement of the NP to the highest, left-most position in the nominal expression, and I will assume the same for Shan. The NP is interpreted in the based-generated position since there is no evidence of this movement being associated with differences in meaning. Thus, the adnominal structure of (17) is proposed to be (26).

(26) ADNOMINAL DISTRIBUTIVE



Just like the adverbial distributive, the adnominal distributive is licensed by a typically overt, plural argument in the sentence. Rather than modifying the VP, the adnominal distributive modifies the classifier phrase. The presented structures parallel the English cases in having two types of adjunct distributives and one argument distributive option. The broad availability of quantifier float leads to the appearance of an extra structure, but this can be treated as a type of adnominal distributive. The base structure of the distributive is the same across the Shan cases except that the adjunct distributives involve elision of the noun.

3. Cross-linguistic representations of distributivity

In sentences that involve pluralities, there is often ambiguity between a cumulative and distributive reading, as in (27). This can be interpreted as the SHARE, *bought one comic book*, applying to each individual child in the KEY *the children*, giving a distributive reading, such as (27a), or it can be interpreted as the SHARE applying to the plurality *the children* as a whole, giving a cumulative reading, as in (27b).

(27) The children bought one comic book.
   a. The children bought one comic book each. DISTRIBUTIVE
   b. The children bough one comic book together. CUMULATIVE

The distributive reading of (27) is accounted for by positing a covert distributive morpheme between the plural subject and the predicate (Champollion, 2016a).

Overt markers of distributivity, like *each* in English (27a), can be found cross-linguistically. For example, there are dependent numerals in Telugu (Balusu, 2006), Kaqchikel (Henderson, 2012), and Bengali (Guha, 2021) as well as distributive morphemes that can appear in multiple syntactic positions such as English *each* and German *jeweils* (Zimmermann, 2002).

Different strategies of marking overt distributivity are associated with different licensing requirements. For example, English *each* requires a semantically plural key. In the adverbial

and adnominal cases, the key must also be syntactically plural, but in the determiner case, it is singular. As discussed in the previous section, the Shan distributive requires a plural key with quantized key and share amounts using overt numeral-classifier expressions for each of these.

Strategies of marking distributivity on the numeral require an overt numeral, as in Telugu and Kaqchikel, but they have different semantic licensing requirements for the key. Kaqchikel requires a plural individual key, while Telugu can make use of an understood plural spatial or temporal key as an alternative to an individual key (Henderson, 2012). Again, the distributive morpheme is acceptable when there is a plural individual key in both Telugu, (28), and Kaqchikel, (29), but if there is no plural individual key in Telugu, distribution to an understood spatial or temporal key is possible, as in (30). This is not available in Kaqchikel as (31) shows.

(28) ii pilla-lu renDu renDu kootu-lu-ni cuus-ee-ru.
these kid-PL two two monkey-PL-ACC see-PAST-3PPL
'These kids each saw two monkeys.' (Telugu, Balusu, 2006: (1), reading (2a))

(29) Rije' x-Ø-ki-chäp el ox-ox kab'.
they COM-A3S-E1P-handle DIR three-RED candy
'They each took three candies.'
FALSE if they took three in total to share.' (Kaqchikel, Henderson, 2012: (408))

(30) Raamu renDu renDu kootu-lu-ni cuus-ee-Du.
Ram two two monkey-PL-ACC see-PAST-3PSG
'Ram saw two monkeys in each location.'
'Ram saw two monkeys in each time interval.' (Telugu, Balusu, 2006: (9)–(10))

(31) *X-e'-in-tz'ët ox-ox batz'.
COM-A3P-E1S-see three-RED monkey
READING SOUGHT: 'I saw three monkeys in each place.'
READING SOUGHT: 'I saw three monkeys each time'
(Kaqchikel, Henderson, 2012: (414))

In the English adnominal distributive, we see that it is not necessary to have an overt singular noun following *each* if there is a plural argument to distribute to, as in (32). If there is no plural argument, it is necessary to specify that the distribution is to times/places, as in (33).

(32) The children/*Ram saw three monkeys each.

(33) Ram saw three monkeys each time/place.

The Shan distributive construction requires an overt numeral and a classifier which matches some properties of the distributive key (properties such as being human, animal, book, etc.). The noun is omitted in cases where the key can be inferred.

These cross-linguistic expressions of distributivity vary in their morphosyntactic representation as well as in their semantic licensing. The overt representation of distributivity seems broadly limited to determiner distributives for the KEY, adverbial distributives that index the KEY, and distributives that attach to a quantity in the SHARE, either by marking it as a dependent indefinite, as in Kaqchikel and Telugu, or as an adnominal distributive.

Semantic accounts of distributivity require universal quantification and specification of min-

imal parts of a whole to distribute to. Common to accounts by Balusu (2006), Henderson (2012, 2014), and Champollion (2016b) is the notion of universal quantification over sub-events where the distributive share specifies the amount associated with the share in each sub-event. This is formally represented in a variety of ways. Balusu (2006) accounts for the Telugu data with universal quantification over sub-events. Henderson (2012, 2014) uses Brasoveanu's (2008) extension of van den Berg's (1996) Dynamic Plural Logic (DPlL). Champollion (2016b) claims that distance distributives are overt representations of D from Link (1983) and Part from Schwarzschild (1996) and implements an account in Neo-Davidsonian event semantics.

The Shan distributive account will follow these but will take advantage of Hamblin semantics for indeterminate pronouns which will be introduced in the next section.

Many accounts of distributive constructions presume distributivity to atomic portions of the KEY (Champollion, 2016b; Henderson, 2021), but the numeral in the Shan [Num Clf *lăj*] construction can be greater than one. This is because the distribution to the key is overtly specified.

The key can be divided into measures of individuals greater than one as shown in (17), (18), and (34). In (34), *săam kô lăj* 'per three students' indicates what number of students each gets two books. Here each group of three gets two books.

(34) lukhén kaw kô tsɤ́ nân, **săam kô lăj** laj pâplik sŏŋ pâp.
student nine CLF.HUM CLF.PL that three CLF.HUM LAJ get book two CLF.BOOK
'Those nine students, each three get two books.'

These are also found in Korean, as in (35). Here, the distributive *-ssik* appears on the distributive share numeral-classifier expression. The distributive key amount is indicated by *twul* 'two'.

(35) Chwuyenca tul-i twul{-ey/tang} noray han-kok-ssik-ul pwul-ess-ta.
performer PL-NOM two song one-CL-ssik-ACC sang
'The performers sang a song per group of two.' (Korean, Choe, 1987: 141)

The account presented in Section 5 will address non-atomic distributivity of this sort.

4. Cross-linguistic accounts of indeterminate pronouns

Examples of indeterminate pronoun uses of *lăj* in Shan can be seen in (36) and (37). A universal quantification interpretation arises with the particle *kɔ*, as in (36), and an NPI interpretation is available with negation, as in (37).

(36) mán-náanŋ tsɔj **mɤ lăj** kɔ tsáaj khám jíntsómkhɔptsăj thŭŋ mán-náaŋ.
3-FEM help when LAJ PRT Jai Kham thankful to 3-FEM
'Whenever she helps Jai Kham, Jai Kham is grateful to her.'

(37) pó pàj hét lɔŋ-mótsăj náj hɤ́n jâw tĕ ʔàm laj kwàa **táaŋ lăj**.
if not.yet do NMLZ-clean in house PFV will NEG can go path LAJ
'If you haven't finished the chores in the house, you cannot go anywhere.'

As (38)–(39) demonstrate, the indeterminate use of *lăj* is not acceptable with the numeral *nɯŋ* 'one'. (38) provides an example of the indeterminate pronoun with universal quantificational force, likely provided by the particle *kɔ*, and negation, provided by *ʔàm* 'not'. (39) is identical except that the numeral *nɯŋ* 'one' is included before the classifier.

(38) lukhén kɔ̂ **lăj** kɔ ʔàm hû khɔtɔ̀p.
student CLF.HUM LAJ PRT NEG know answer
'Each student did not know the answer.'

(39) #lukhén nɯŋ kɔ̂ **lăj** kɔ ʔàm hû khɔtɔ̀p.
student one CLF.HUM LAJ PRT NEG know answer
intended: 'Each student did not know the answer.'

The meaning of (38) suggests a strong connection to distributivity, which will be reflected in the analysis, but there is a clear structural difference between the distributive use and the indeterminate use. The universally quantified indeterminate use is clearly identified by the particle *kɔ* that appears between the subject and the predicate in (38), and the distributive use requires numerals to describe the distributive key and distributive share amounts.

In (40) are examples of indeterminate pronouns in Japanese. (40a) has a universal interpretation, and (40b) is a *wh*-question. In Japanese, particular particles are associated with specific interpretations, *-mo* for universal and *ka* for questions. In Shan, *wh*-questions are the unmarked form, and the particle *kɔ* is associated with a universal interpretation.

(40) a. [[**Dono hon**-o yonda] kodomo] -**mo** yoku nemutta.
which book-ACC read child -MO well slept
'For every book x, the child who read x slept well.'
b. Taro-wa [[**dare**-ga katta] mochi]-o tabemasita **ka**?
Taro-TOP who-NOM bought rice cake-ACC ate Q
'Who is the x such that Taro ate rice cakes that x bought?'
(Kratzer and Shimoyama, 2002: (2))

Kratzer and Shimoyama (2002) propose a unified account of indeterminate pronouns based on a Hamblin semantics (Hamblin, 1973) for questions. In (40), we see two interpretations of indeterminate pronouns in Japanese. In Hamblin semantics, an expression is interpreted as the set of its typical meaning. Question words like *who* or *which book* are interpreted as sets of alternatives that could serve as an answer to the question. These can combine pointwise through Hamblin Funcational Application, shown in (41), until they meet an operator that quantifies over them. For example, *dare* 'who' is interpreted as a set of humans, as in (42a), the predicate *nemutta* is the set containing the function from individuals to functions from words to truth values, as in (42b), and the question *dare nemutta* 'who slept?' is interpreted as the set of propositions where someone slept, as in (42c).

(41) **Hamblin Functional Application** (Kratzer and Shimoyama, 2002)
If α is a branching node with daughters β and γ, and $[\![\beta]\!]^{w,g} \subseteq D_\sigma$ and $[\![\gamma]\!]^{w,g} \subseteq D_{\langle\sigma,\tau\rangle}$, then $[\![\alpha]\!]^{w,g} = \{a \in D_\tau : \exists b \exists c[b \in [\![\beta]\!]^{w,g} \,\&\, c \in [\![\gamma]\!]^{w,g} \,\&\, a = c(b)]\}$.

(42) a. $[\![\text{dare}]\!]^{w,g} = \{x : \mathsf{human}(x)(w)\}$
b. $[\![\text{nemutta}]\!]^{w,g} = \{\lambda x \lambda w'.\mathsf{slept}(x)(w')\}$
c. $[\![\text{dare nemutta}]\!]^{w,g} = \{p : \exists x[\mathsf{human}(x)(w) \,\&\, p = \lambda w'.\mathsf{slept}(x)(w')]\}$

Shimoyama (2006) and Yatsushiro (2009) analyze Japanese distributive *-mo* as a universal quantifier. For example in (43), the indeterminate pronoun construction *dono gakusei* 'which student' would be interpreted as a set of students. The semantics for *-mo* from Shimoyama (2006) is given in (44), and the truth conditions of (43) is shown in (45). This is similar to

a Hamblinized univeral quantifier, except that universal quantification is over the alternatives expressed by the prejacent.

(43) Dono gakusei-mo odotta.
which student-MO danced
'Every student danced.' (Shimoyama, 2006: (25a))

(44) For $[\![\alpha]\!]^g \subseteq D_e$,
$[\![\alpha \textbf{ mo}]\!]^g = \{\lambda P.\forall x[x \in [\![\alpha]\!]^g \rightarrow P(x) = 1]\}$ (Shimoyama, 2006: (27))

(45) Truth conditions for (43):
$\{\forall x[x \in \{y\colon \mathsf{student}(y)\} \rightarrow \mathsf{dance}(x)]\}$ (Shimoyama, 2006: (28a))

This account can easily be extended to indeterminate pronouns in Shan. The particle *kɔ* appears to be a good candidate for the Shan version of *mo*. In this way we can account for (38):

(46) Truth conditions for (38):
$\{\forall x[x \in \{y\colon \mathsf{student}(y)\} \rightarrow \neg\mathsf{know}(x, \mathsf{answer})]\}$
For every x in the alternative set of students, it is false that x knows the answer.

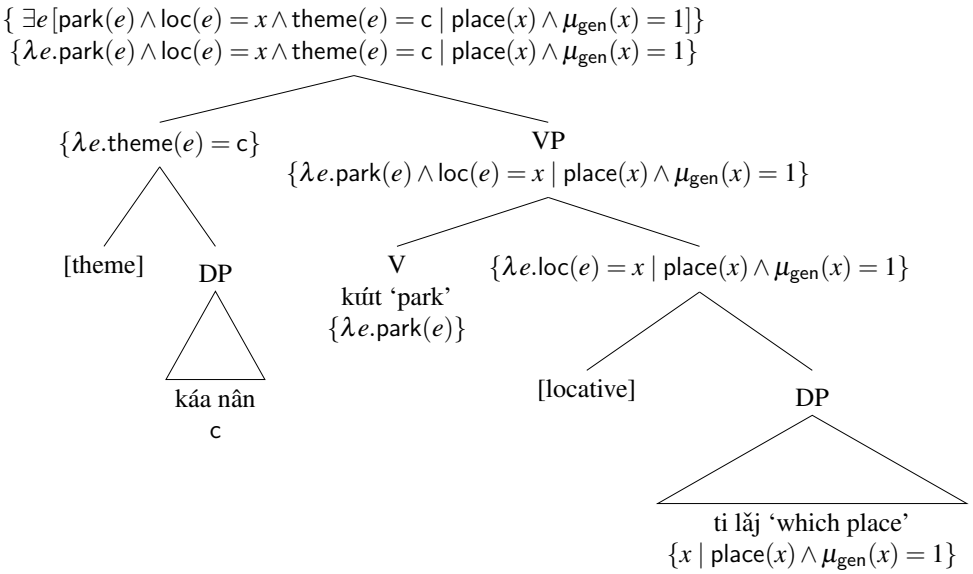
The account below will reformulate the Hamblin semantics account of Kratzer and Shimoyama (2002) in Neo-Davidsonian algebraic event semantics (Davidson, 1966), following Champollion (2016b). Event semantics will be necessary to model adnomial distributivity but will be used throughout for a unified account. This will make use of θ heads and type-shifters for definite/indefinite expressions from Champollion (2016b) as in (47), and the property of events will be existentially closed at the top of the clause. Predicates of events will combine using Predicate Modification.

(47) a. Type shifter for definites: $\lambda\theta\lambda x\lambda e.\theta(e) = x$
b. Type shifter for indefinites: $\lambda\theta\lambda P\lambda e.\exists x[P(x) \wedge \theta(e) = x]$
(slightly adjusted from Champollion (2016b: (76)) for readability)

(48) models a compositional account of the *wh*-question in (5), repeated below. The indeterminate expression *ti lăj* 'which place' is interpreted as a set of alternative locations. The result of composition is a set of propositions indicating there was an event of 'that car' being parked that vary on the location of the parking. I will assume that a locative θ head can be introduced in this construction in the position shown, and the DP *ti lăj* 'which place' is a definite expression. In (48), I have shown existential closure of the event as a separate step, but in future derivations it will be assumed to take place after arguments have been introduced. Shan is a GENERAL NUMBER language in the sense of Corbett (2000), so in the semantics, properties of individuals or events are assumed to apply to both atoms and sums of atoms, except for measure expressions that will either apply to atoms or specific measures of sums.

(5) káa nân kʉ́t ti **lăj**?
car that park place LAJ
'**Where** (lit. **which** place) is that car parked?'

(48) **Wh-question**



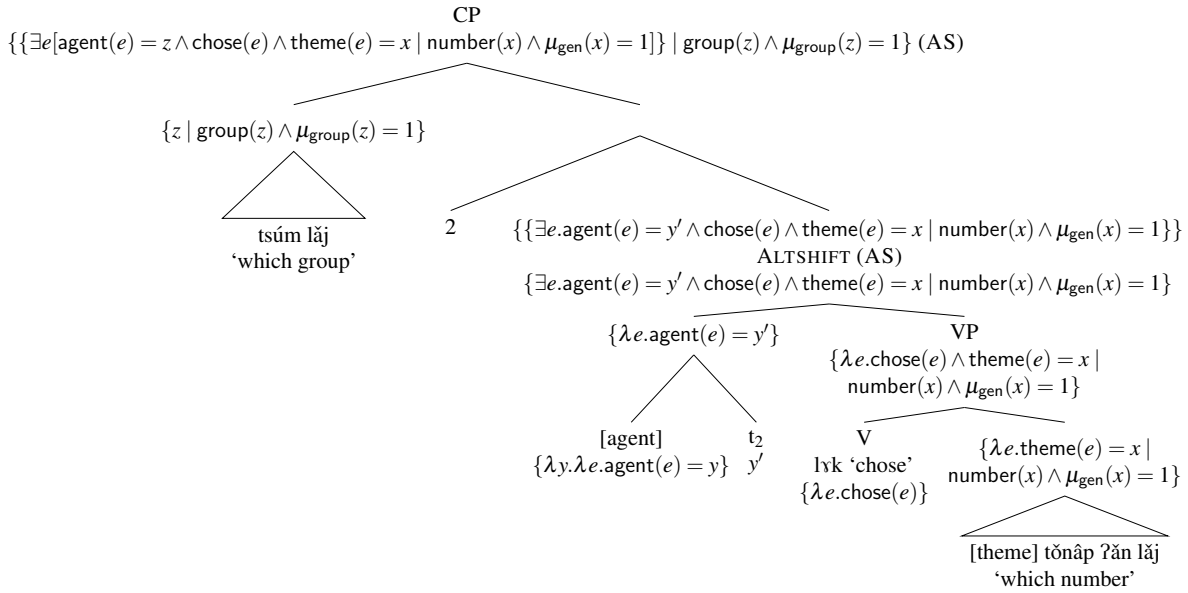
If there are multiple wh-questions in a sentence, a pair-list response is possible as in (10) which can be answered by describing the groups ('the first group', 'the left group') and listing the numbers that they chose. (49) also allows a pair-list answer. It is not obligatory to mark a morphological distinction between 'which number' and 'which numbers', and there does not appear to be a uniqueness effect for *lăj*-questions. This distinction is relevant in the literature on questions (e.g., Dayal (1996)), but will not be addressed further here.

(10) tsúm **lăj** lɤk tŏnâp ʔăn **lăj**?
group LAJ choose number CLF.G LAJ
'Which group chose which number(s)?'
(Context: 3 groups of students play a game. Each group picks two numbers (1–9).)

(49) khúsɔ̆n kô **lăj** tsɔjthɛ̆m lukhén kô **lăj** làa?
teacher CLF.HUM LAJ help student CLF.HUM LAJ Q.PRT
'Which teacher helped which student(s)?'
(Context with 3 teachers; Answer pointing to students: 'Teacher Lao Saeng helped these students, Teacher Mo Horm helped these, Teacher Lao Nguen helped these…')

Multiple-wh questions with pair-list answers require some extra machinery. See Fox (2012); Abels and Dayal (2023); Xiang (2023); Roelofsen and Dotlačil (2023) for in-depth discussion. To generate a family-of-questions (FoQ) for multiple wh-questions, we can use ALTSHIFT from Kotek (2016). ALTSHIFT is an operator that takes as input alternative propositions and shifts them into the ordinary domain. At the node where ALTSHIFT first applies in (50), it shifts the focus alternative propositions to an ordinary set of propositions, the focused value of which is a set of sets of propositions. This way, possible answers to one question word exist as a set of propositional alternatives for each possible answer to the higher question word.

(50) **Multiple wh-questions**



Thus, (10) in the specified context gives the family-of-questions $\mathbb{Q}$ in (51), assuming that there are three groups (A, B, C) choosing numbers between 1 and 9.

(51) For example (10), $\mathbb{Q}=\left\{\left\{\begin{array}{c}\text{A chose 1,}\\\text{A chose 2,}\\\text{A chose 3,}\\\dots\end{array}\right\},\left\{\begin{array}{c}\text{B chose 1,}\\\text{B chose 2,}\\\text{B chose 3,}\\\dots\end{array}\right\},\left\{\begin{array}{c}\text{C chose 1,}\\\text{C chose 2,}\\\text{C chose 3,}\\\dots\end{array}\right\}\right\}$

Xiang (2023) argued convincingly that wh-questions are not obligatorily exhaustive, so it is possible that not every member of the family-of-questions would need to be answered. To account for this, I propose to flatten the family-of-questions to include (non-)exhaustive alternatives, following Xiang's (2023) observations, by using (52). This generates a set of propositions that each completely answer a sub-set of the sub-questions in the family-of-questions $\mathbb{Q}$. Applying **CCA** to (51) could look like (53) in a context where each group chose two numbers. Then, the complete answer to the questions would be the unique, maximally informative true proposition.

(52) **Combine consistent alternatives**:
$\mathbf{CCA}(\mathbb{Q}):=\{p\mid\exists\mathbb{Q}'_{(stt)t}\subseteq\mathbb{Q}_{(stt)t}[\forall Q_{(st)t}\in\mathbb{Q}'_{(stt)t}[\exists q\in Q_{(st)t}[\exists w[q(w)\wedge p(w)]]]]\}$
The set of propositions p such that there is a subset of the family-of-questions $\mathbb{Q}$ where every sub-question contains a proposition q such that some q worlds are p worlds.

(53) $\mathbf{CCA}(\mathbb{Q})=\left\{\begin{array}{c}\text{A chose 1+2 and B chose 3+4,}\\\text{A chose 1+2 and B chose 3+4 and C chose 5+6,}\\\text{A chose 1+2 and B chose 3+4 and C chose 6+7, }\dots\end{array}\right\}$

In the following section, this account of indeterminate pronouns will be extended to account for distributive uses of *lăj*.

5. Hamblin semantics for *lăj* in distributive constructions

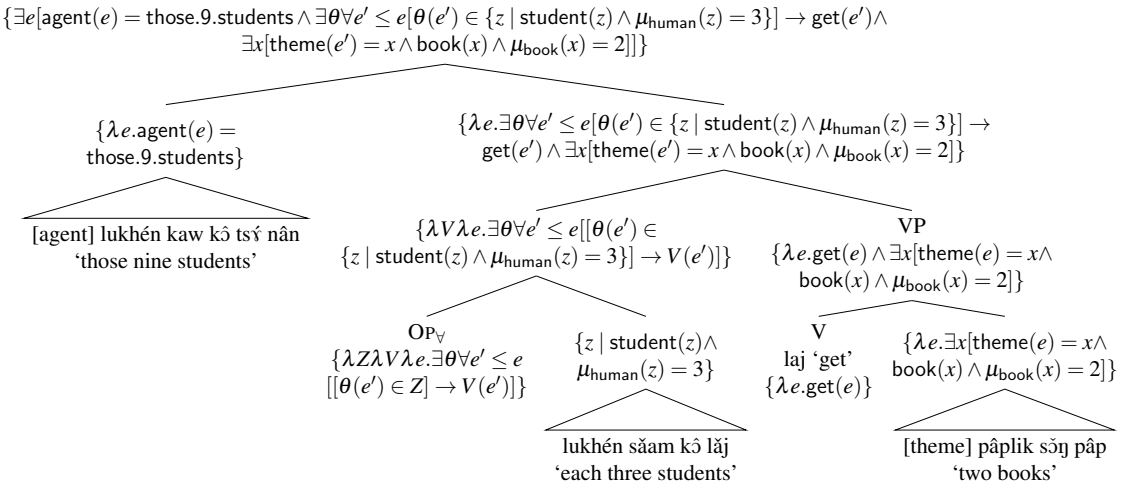
The generalized quantifier from Kratzer and Shimoyama (2002) associates two properties to the same individual while the adnominal distributive connects two distinct entities, so we modify the quantifier for distributive cases. We can connect the distributive key and distributive share together using the event that applies to both of them following Champollion (2016b). This motivates using Neo-Davidsonian event semantics. Distributing over properties of events is already useful to account for examples like (21), which distribute over times or places.

The universal quantifier semantics used for *kɔ* can be adapted to represent the distributive operator in a Neo-Davidsonian framework, as in (54). The predicate property is now a property of events, type $\langle s,t\rangle$, and the operator quantifies over events such that there is a theta role associated with that event that maps to a member of the alternative set that the distributive operator first combines with. This fits the licensing requirement of this operator for a plurality to have already been associated with a theta role of the complete sentence event, as in the adnominal and adverbial cases. The derivation for adverbial (34) is given in (55).[7]

(54) For $[\![\alpha]\!]^g \subseteq D_e$,
$[\![\mathbf{Op}_\forall\ \alpha]\!]^g = \{\lambda V_{\langle s,t\rangle}\lambda e.\exists\theta\forall e' \leq e[[\theta(e') \in [\![\alpha]\!]^g] \rightarrow V(e')]\}$

(34) lukhén kaw kô tsɤ́ nân, **săam kô lăj** laj pâplik sɔŋ pâp.
student nine CLF.HUM CLF.PL that three CLF.HUM LAJ get book two CLF.BOOK
'Those nine students, each three get two books.'

(55) **Adverbial distributive**



An important factor when applying universal quantification to non-atomic properties is that it is necessary to rule out overlap. This will not be formally implemented here, but would be possible to do using, for example, a partition (a non-overlapping cover) from Ionin and Matushansky (2006). The alternatives described by the distributive *săam kô lăj* are non-overlapping parts of their antecedent, *lukhén kaw kô tsɤ́ nân* 'those nine students'.

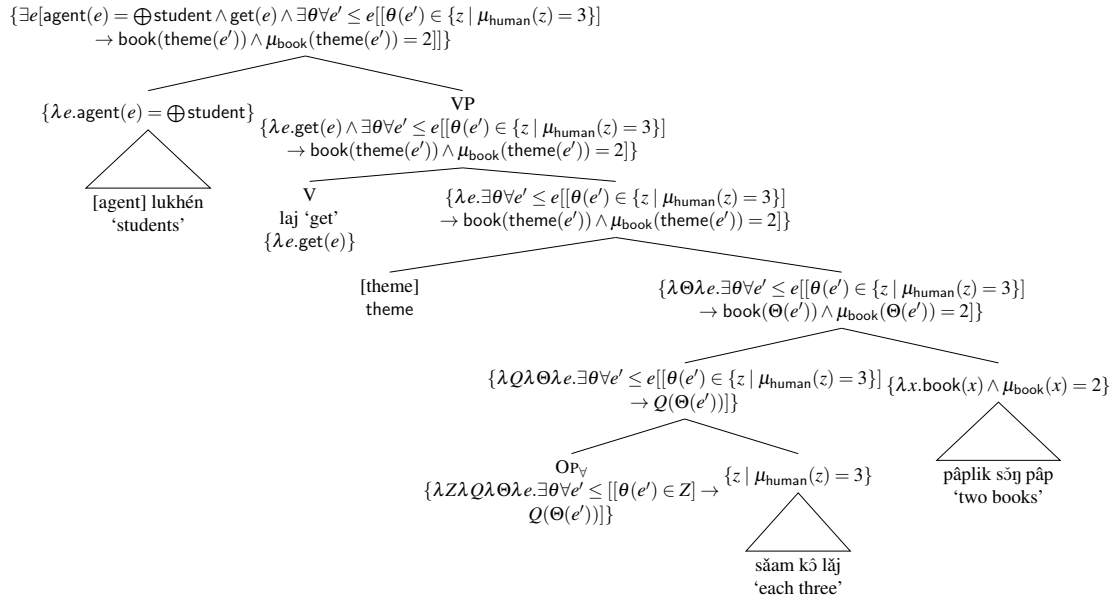
(56) shows the composition for the adnominal example in (17), repeated below. Following Champollion (2016b), the adnominal distributive operator takes in the theme theta role as an

[7] The semantics of the anaphoric plural demonstrative expression 'those nine students' is abbreviated for space.

argument. The distributive key amount is anaphoric to the subject, *lukhén* 'students'. The distributive phrase combines with the property of the argument it modifies and its theta role.

(17) lukhén laj pâplik **săam kô** **lăj** sŏŋ pâp.
student get book three CLF.HUM LAJ two CLF.BOOK
Each group of three students gets two books.

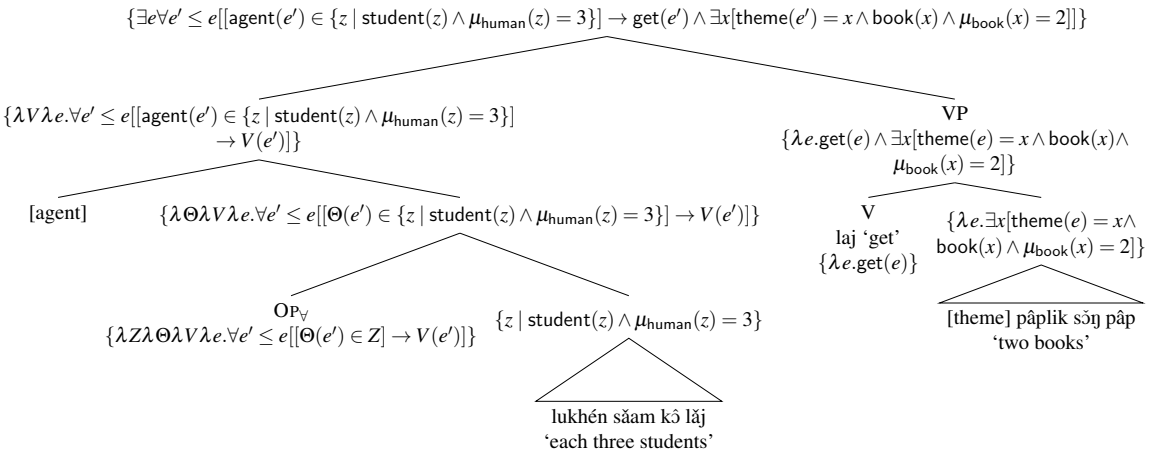
(56) **Adnominal distributive**



For a determiner distributive like (18), repeated below, the semantics would differ from (54) only in taking a theta role as an argument rather than relying on a previously introduced antecedent expression in the clause. When the distributive phrase is part of a theta-role bearing expression, it takes that theta role as a semantic argument. (57) shows the composition of (18).

(18) lukhén **săam kô** **lăj** laj pâplik sŏŋ pâp
student three CLF.HUM LAJ get book two CLF.BOOK
'Each three students get two books.'

(57) **Determiner distributive**



This has demonstrated that the Hamblin-style indeterminate pronoun account can be unified with an account of distributivity with minimal adjustments. A key difference between the account from Champollion (2016b) and the one presented here is that the distributive operator in this account applies to alternatives. The Shan distributive construction is relatively similar to adverbial and adnominal *each* as analyzed by Champollion (2016b). The main difference empirically is that the distribution over the key is to atomic parts in English but does not have to be in Shan. This is because the key distribution is overtly specified.

6. Conclusion

This paper presents novel data from Shan, an understudied language, on expressions of distributivity using question word *lǎj* 'which'. The same morpheme that functions as the indeterminate pronoun used to express the question word 'which' is also used to mark overt distributivity. This distributive construction is unique in this overlap between the distributive and indeterminate pronoun constructions as well as in the fact that the distributive key amount is overtly specified, leading to distribution to non-atomic keys. A Hamblin semantics of indeterminate pronouns with a covert distributivity operator accounts for this data. Additionally, a new method of flattening family-of-questions has been proposed that allows for non-exhaustive interpretations of multiple-wh questions. This approach easily deals with cases of non-atomic distributive keys, which is relevant for distributivity cross-linguistically.

References

Abels, K. and V. Dayal (2023). On the syntax of multiple sluicing and what it tells us about *wh*-scope taking. *Linguistic Inquiry 54*(3), 429–477.

Aikhenvald, A. Y. (2000). *Classifiers: A Typology of Noun Categorization Devices*. Oxford Studies in Typology and Linguistic Theory. Oxford: Oxford University Press.

Balusu, R. (2006). Distributive reduplication in Telegu. In C. Davis, A. R. Deal, and Y. Zabbal (Eds.), *Proceedings of North East Linguistic Society (NELS) 36*, Volume 1, Amherst, MA, pp. 39–52. GLSA.

Brasoveanu, A. (2008). Donkey pluralities: Plural information states versus non-atomic individuals. *Linguistics and Philosophy 31*(2), 129–209.

Chaiphet, K. (2023). *The Structure of Discontinuous Noun Phrases in Thai: Right-dislocation and Quantifier Float*. Ph. D. thesis, Stony Brook University, Stony Brook, NY.

Champollion, L. (2016a). Covert distributivity in algebraic event semantics. *Semantics and Pragmatics 9*(15), 1–65.

Champollion, L. (2016b). Overt distributivity in algebraic event semantics. *Semantics and Pragmatics 9*(16), 1–65.

Champollion, L. (2019). Distributivity in formal semantics. *Annual Review of Linguistics 5*, 289–308.

Choe, J.-W. (1987). *Anti-Quantifiers and a Theory of Distributivity*. Ph. D. thesis, University of Massachusetts Amherst, Amherst, MA.

Corbett, G. G. (2000). *Number*. Cambridge Textbooks in Linguistics. Cambridge, UK: Cambridge University Press.

Davidson, D. (1966). The logical form of action sentences. In N. Rescher (Ed.), *The Logic of Decision and Action*. Pittsburgh: University of Pittsburgh Press.

Dayal, V. (1996). *Locality in Wh-Quantification: Questions and Relative Clauses in Hindi*.

Dordrecht: Kluwer.
Ethnologue (2021). Shan. *Ethnologue: Languages of the World*.
Fox, D. (2012). Multiple wh-questions: Uniqueness, pair-list and second order questions.
Guha, I. (2021). Dependent numerals in Bengali: A case for covert adverbial D-operators. In P. Grosz, L. Martí, H. Pearson, Y. Sudo, and S. Zobel (Eds.), *Proceedings of Sinn Und Bedeutung 25*, pp. 376–393.
Hamblin, C. L. (1973). Questions in Montague English. *Foundations of Language 10*(1), 41–53.
Henderson, R. (2012). *Ways of Pluralizing Events*. Ph. D. thesis, University of California Santa Cruz, Santa Cruz, CA.
Henderson, R. (2014). Dependent indefinites and their post-suppositions. *Semantics and Pragmatics 7*(6), 1–58.
Henderson, R. (2021). Dependent numerals in Kaqchikel. In P. Cabredo Hofherr and J. Doetjes (Eds.), *The Oxford Handbook of Grammatical Number*. Oxford University Press.
Ionin, T. and O. Matushansky (2006). The composition of complex cardinals. *Journal of Semantics 23*(4), 315–360.
Jenks, P. (2011). *The Hidden Structure of Thai Noun Phrases*. Ph. D. thesis, Harvard University, Cambridge, MA.
Kotek, H. (2016). On the semantics of wh-questions. In N. Bade, P. Berezovskaya, and A. Schöller (Eds.), *Proceedings of Sinn Und Bedeutung 20*, pp. 430–447.
Kratzer, A. and J. Shimoyama (2002). Indeterminate pronouns. The view from Japanese. In Y. Otsu (Ed.), *The 3rd Tokyo Conference on Psycholinguistics*, pp. 1–25. Hituzi Syobo.
Kuroda, S.-Y. (1965). *Generative Grammatical Studies in the Japanese Language.* Ph. D. thesis, MIT, Cambridge, MA.
Link, G. (1983). The logical analysis of plurals and mass terms: A lattice-theoretical approach. In R. Bäuerle, C. Schwarze, and A. von Stechow (Eds.), *Meaning, Use and Interpretation of Language*, pp. 127–145. Berlin: de Gruyer.
Moroney, M. (2021). *Definiteness and Quantification: Evidence from Shan*. Ph. D. thesis, Cornell University, Ithaca, NY.
Roelofsen, F. and J. Dotlačil (2023). Wh-questions in dynamic inquisitive semantics. *Theoretical Linguistics 49*(1-2), 1–91.
Schwarzschild, R. (1996). *Pluralities*. Number 61 in Studies in Linguistics and Philosophy (SLAP). Dordrecht and Boston: Kluwer Academic Publishers.
Shimoyama, J. (2006, June). Indeterminate Phrase Quantification in Japanese. *Natural Language Semantics 14*(2), 139–173.
Simpson, A. (2005). Classifiers and DP structure in Southeast Asia. *The Oxford handbook of comparative syntax*, 806–838.
van den Berg, M. (1996). *Some Aspects of the Internal Structure of Discourse*. Ph. D. thesis, University of Amsterdam, Amsterdam.
Xiang, Y. (2023). Quantifying into wh-dependencies: Multiple-wh questions and questions with a quantifier. *Linguistics and Philosophy 46*(3), 429–482.
Yatsushiro, K. (2009). The distribution of quantificational suffixes in Japanese. *Natural Language Semantics 17*(2), 141–173.
Zimmermann, M. (2002). *Boys Buying Two Sausages Each: On the Syntax and Semantics of Distance-Distributivity*. Ph. D. thesis, University of Amsterdam, Amsterdam, Netherlands.