https://doi.org/10.1016/j.lingua.2023.103571

Definiteness in plural reference in Shaoxing Wu

Jennifer Yao [a], Stephen Matthews [b]

[a] School of Professional Education and Executive Development, The Hong Kong Polytechnic University, Hong Kong

[b] Department of Linguistics, The University of Hong Kong, Pokfulam, Hong Kong

Abstract: This study reports a previously undescribed strategy for definite plural reference in the Wu dialect of Shaoxing: lian^{22}-CL-NP, where 'lian^{22}' together with a classifier marks a plural reference. We examine in depth its behaviours in different definite environments. We then show that the 'lian^{22}-CL-NP' structure is predominantly used to represent a definite plural in typical definite contexts, except for part-whole bridging situations. Moreover, we argue that the numeral-classifier 'lian (two)-CL' in Shaoxing Wu undergoes a process of grammaticalisation wherein the lexical meaning of quantity 'two' is lost and the 'lian-CL' constitutes a single tonal unit which is employed to encode the meaning of 'a few'. Finally, we discuss several possible ways to derive the syntactic structure of DP with numeral and classifier elements. In doing so, we illustrate that all of the above properties of the 'lian^{22}-CL-NP' can be captured by the two-head analysis, in which the number and classifier are assumed to instantiate two distinct heads.

Keywords: Chinese DPs; Definiteness; Lian^{22}-CL-NP; Plural strategy; Shaoxing Wu

1. Introduction

Research on definiteness in Chinese languages has established that a range of options are used among different dialects. In addition to bare NPs and demonstratives, a classifier-noun (CL-NP) sequence without an accompanying numeral, which is called the 'bare-classifier' pattern, has been reported to represent definite reference, especially in southeastern Chinese languages. For example, the bare-classifier-noun structure is predominantly used to represent definite noun phrases in Cantonese and is normally acceptable in Wu dialects[1], while it sporadically appears in other Chinese dialects (e.g. Matthews and Yip, 1994; Cheng and Sybesma, 1999, 2005; Li and Bisang, 2012; Simpson, 2017; Yao, 2018). The following are examples of definite CL-NP in Chinese dialects.

Cantonese[2]

(1)***(Zek) gau** soeng gwo maalou (Cheng and Sybesma, 1999).
CL dog want cross road.
'The dog wants to cross the road'.

Wenzhou Wu

[1] Wu Chinese is a major group in Chinese that is situated between Mandarin and Min-Cantonese groups. It is spoken in the Yangtze Delta area, including Shanghai, Zhejiang Province and southern Jiangshu Province. Furthermore, it is divided into six subgroups. Roughly, Wu dialects are also classified into two subtypes: southern Wu and northern Wu, with the latter being normally viewed as the most typical representative as the topic prominent language in Chinese with a predominant SOV word order. Both Jinyun Wu, located in the Jin-Qu area and Wenzhou Wu as mentioned in this paper belong to the southern Wu, thus sharing some commons with Min Chinese. Meanwhile, Shaoxing Wu, which is discussed in this paper, is a typical member of northern Wu. It is spoken in and around the city of Shaoxing and has approximately 5,270,977 speakers.

[2] The original versions of the cross-language data cited from other works related to the current topic are retained in this paper. Cantonese data cited from Cheng and Sybesma (1999) are recorded with the Jyutping system, and the Wenzhou Wu data from Cheng and Sybesma (2005) are transcribed with IPA without tones. Several Jinyun Wu data from Simpson (2017) are also transcribed with IPA without tones. However, all Shaoxing Wu data presented in this paper are recorded with IPA system with tones. The data in Mandarin are recorded with Pinyin.

(2) **dɣu kau** i tsau-ku ka-løy (Cheng and Sybesma, 2005).
CL dog want walk-cross road
'The dog wants to cross the road'.

Shaoxing Wu

(3) 儂 **塊** **地** 板 揩-揩 伊 清 爽 (Yao, 2018).
no^{ʔ2} **khue^{33} di^{11}** pɛ̃^{13} kha^{1}- ^{1}kha^{33} i^{11} tɕhiŋ^{53} saŋ^{11}
you CL floor wipe-wipe it clean
'Please clean **the floor'.**

Jiyun Wu

(4) nɛdʑiɔdʑiəmdɔ'o jiəmman ma. **gø dusuguɔ** dø ma (Simpson, 2017)
USC famous very CL library big very
'USC is very famous. **The library** is very big'.

Differing from Cantonese, the bare CL-NP structure in Wu dialects alternates with the bare noun phrase to refer to a definite individual/entity in many cases (Yao, 2018). Simpson (2017) described such an alternation in Jinyun Wu in relation to the four different contexts for definites initially proposed by Hawkins (1978). These contexts include anaphoric definite, bridging, reference to salient objects visible or nonvisible and with or without personal relations with the speaker, and specifically unique or directly connected to the speaker. Simpson further argued that no free alternation occurs between bare NP and bare CL-NP patterns in Jinyun Wu. These two options highlight different aspects of definiteness, showing strong parallels to languages with strong and weak definite articles.

However, the definite referents in all cases discussed by Simpson (2017) are singular and countable. As such, cases with plural and mass definite reference have not yet been reported. Will the same strategy be employed to encode plural definite reference as in cases with singular referents? The answer seems negative, at least in Shaoxing Wu. Empirically, Shaoxing Wu adopts another undescribed definiteness option, 'liaŋ^{22}-CL-NP', instead of bare CL-NP to represent plural definite reference. This option is exemplified in (5), where the first 'liaŋ^{22}-gəʔ ŋiŋ^{23}' encoding indefinite 'some workers' is employed and occupies the postverbal position. Meanwhile, the second one in the preverbal position refers back to the referents introduced by the first 'liaŋ^{22}-gəʔ ŋiŋ^{23}' and is understood as definite.

(5) tso^{33} vəʔ^{23} dʑi^{231} maʔ ŋo^{113} tɕʰiŋ^{335}-ləʔ **liaŋ^{22}gəʔ ŋiŋ^{23}**. pieʔ^{41}ku^{11}, **liaŋ^{22}-gəʔ ŋiŋ^{23}**
do not finish Prt I hire-PFV a few-CL person. However, a few-CL person

tu^{33} vəʔ^{23} naʔ^{23}geʔ^{23} ɦuᴇ^{41} tso^{33} la.
all not very can do Prt
'I cannot finish everything myself alone, so I hired **some workers** to help me. However, **the workers** are not helpful'.

The interpretation of the second 'liaŋ^{22}-gəʔ ŋiŋ^{23}' in the above data (5) depends on that of a preceding expression, which is a typical anaphoric use of a definite. We then need to examine other definite contexts where plural reference is involved. Will the 'liaŋ^{22}-CL-NP' structure be employed in all definite environments, or will it alternate with other patterns, such as bare NP, CL-NP, Num-CL-NP and Dem-CL-NP, to represent different aspects of definiteness in Shaoxing Wu? The varieties of definiteness marking in Chinese languages have received a substantial amount of attention in the literature within different theoretical frameworks (e.g. Cheng & Sybesma, 1999; Chen, 2009; Jenks, 2018; Dayal and Jiang, 2021; Bremmers et al., 2021). However, few attempts have been made to

distinguish singular definite reference from plural reference in Chinese. Thus, a limited number of studies have focused on definiteness in plural reference. New data from Shaoxing Wu show that it employs distinctive strategies to represent singular and plural definite reference. Unlike singular definite reference, which is dominantly encoded by the bare CL-NP, Shaoxing Wu prefers the expression 'lianη^{22}-CL-NP' to denote plural definite reference.

This study aims to describe this novel definiteness strategy for plural reference in Shaoxing Wu. The work examines the behaviours of this plural definite reference in different types of definite contexts as proposed by Simpson (2017) as well as its alternation with other options, such as bare nouns, bare CL-NPs and Dem-CL-NPs. Furthermore, the data presented in this paper indicate that the numeral lianη in Shaoxing Wu undergoes a process of grammaticalisation: lianη 'two' > lianη + classifier 'a few', wherein the lexical meaning quantity 'two' is bleached and 'lianη-CL' is employed with noun phrases to represent (in)definite plural reference. The paper is organised as follows. Part 2 provides a brief review of the two-way distinction between strong versus weak definites, which is parallel to the traditional approaches of familiarity versus uniqueness to the concept of definiteness in the literature. This part also presents an introduction of the types of definite environments with which the lianη^{22}-CL-NP structure interacts. Next, part 3 is an introduction of 'lianη' and 'lianη^{22}-CL-NP' in Shaoxing Wu. Then, the definiteness properties of 'lianη^{22}-CL-NP' as well as its alternations with other definiteness strategies will be discussed in part 4. In part 5, we discuss the 'lianη^{22}-CL' as a grammaticalisation of plural marking strategy in Shaoxing Wu. Moreover, we provide a structural analysis of the approximative 'lianη^{22}-CL-NP' construction where the 'lianη^{22}-CL' forms a tone unit and is motivated to move up to the D^0 position to obtain a(n) (in)definite reading in part 6. The whole picture of the definiteness strategies in Shaoxing Wu, including its distinct means to represent singular definite referents and a comparison with other Chinese languages, such as Cantonese and Mandarin, will be touched upon in future work.

2. Definiteness and types of definite environments

2.1 What is definite?

How do we define the notion of definiteness? In the literature, rich, descriptive and theoretical work have been conducted on definiteness and related issues in natural languages. In addition, several attempts have been made to characterise definiteness, especially within formal approaches. Mainly, two terms are used: uniqueness and familiarity, which are successively proposed to be the central notions underlying definite reference (Russell, 1905; Abbott, 2004; Christophersen, 1939; Heim, 1982; Kamp, 1981 and many others). However, definite descriptions across languages sometimes favour familiarity over uniqueness, whereas the reverse applies in other cases. Lyons (1999) and Schwarz (2009) argued that uniqueness and familiarity are central notions to license the use of definite expressions; moreover, these two concepts are encoded by distinct forms in languages, such as weak versus strong definites in German. According to Schwarz, the distribution of strong and weak definites in German shows strong parallels to the distinction of familiarity versus uniqueness. Strong definites are predominantly used in anaphoric cases, embracing the character of familiarity, while weak definites primarily appear in immediate- and larger-situation definite contexts in Hawkins' terms, indicating the characteristics of uniqueness. For the associative (or bridging) uses that Hawkins discussed, Schwarz argued that some qualify as strong and others as weak. Key examples from Schwarz (2009) are given in (6) and (7).

Strong definite

(6) In der New Yorker Bibliothek gibt es ein Buch über Topinambur. Neulich

in the New York library exists there a book about topinambur recently

war ich dort und habe #**im /in dem Buch** nach einer Antwort auf die Frage
was I there and have #in.the/in the book for an answer to the question
gesucht, ob man Topinambur grillen kann.
searched if one topinambur grill can
'In the New York Public Library, there is a book about topinambur. Recently, I was there and searched **in the book** for an answer to the question of whether one can grill topinambur'.
(Schwarz 2009:30)

Weak definite

(7) Der Empfang wurde **vom /#von dem Bürgermeister** eröffnet.
the reception was by.the /#by the mayor opened
'The reception was opened **by the mayor**'. (Schwarz 2009:40)

Following Schwarz (2009), several studies have argued that the strong versus weak distinction creates underlying definiteness paradigms in typologically diverse languages with or without articles, such as bare NPs versus demonstratives distinctions in Mandarin (Jenks, 2018; Dayal and Jiang, 2021; Bremmers et al., 2021), bare NPs versus bare CL-NPs in Jinyun Wu (Simpson, 2017) and many other languages (Aguilar-Guevaraet al., 2019). In this study, we have no intention to argue that the uniqueness approach is better than the familiarity one or vice versa. Instead, we concentrate attention on the marking strategies adopted in different definite environments. In the current work, an entity or individual denoted by the nominal phrase is assumed to be definite as long as it is globally unique or its reference has already been established directly or indirectly (inferably) in the previous context.

2.2 Types of definite environments

Hawkins (1978) identified eight different usage types of the definite article in English, which have been emended and grouped into four major classes of definite environments, as shown below:

a. Anaphoric use

(8) Mary bought a book. The book was expensive.

b. Immediate situation

(9) Beware of the dog!

c. Larger situation

(10) The sun rises every morning.

d. Bridging (Clark, 1975, Hawkins' associative anaphora)

(11) Mary bought a book. The author is very famous. (producer-product)

(12) John was driving down the street. The steering wheel was cold. (part-whole)

The so-called larger situation above, also known as global uniqueness, involves reference to entities or individuals that uniquely meet the descriptive content of the definite description according to world knowledge, such as 'the sun' and 'the earth'. Given that this study focuses on plural definite reference, the larger situation involving singular reference will not be discussed. In addition, the immediate situation is elaborated as the contexts with referents that are (non)visible and with(out) contrast to distinguish the demonstrative from the definite in Shaoxing Wu. According to Hawkins, the entities or individuals in the immediate situation are specific and unique only for the parties involved in the current context but not in all situations. This type of definite use may be separated into several subtypes, such as recognitional use and deictic use (Himmelmann, 1996). Recognitional referents represent another type of referent that is identifiable by both the speaker and the listener because of

their previous shared experience or specific knowledge, without being an element of the current discourse universe. These uses in Shaoxing Wu are exemplified in the following sections, as in (30)-(31), (35)-(36) and (57)-(58). They may overlap with two other types of definite contexts, namely, bridging referents and referents involving close personal relations. Both the structures of CL-NP and bare NP may be employed to represent recognitional referents in Shaoxing Wu, and the choice between these two options is regulated by the degree of specificity. Thus, we do not discuss recognitional referents separately[3]. Building on Simpson (2017), this study will examine the plural definite reference represented by liaŋ^{22}-CL-NP in four definite environments: (1) anaphoric context, (2) bridging or associative anaphora context, (3) context with referents (non)visible and with(out) contrast and (4) context with (non) speaker's kin and personal relations.

3. Liaŋ (two) and liaŋ^{22}-CL-NPs in Shaoxing Wu

3.1 Two variants of liaŋ in Shaoxing Wu

The word liaŋ (two) in Shaoxing Wu has multiple usages which are distinguished by tones. One is *liaŋ^{113}* with a rising tone, which means 'two', and the other is *liaŋ^{22}* with a level tone, serving as an approximation which means 'a couple of' or 'a few'.

Just as Mandarin has two words for 'two': *er^{51}* (二) and *liang^{214}* (两), *liaŋ^{113}* in Shaoxing Wu also has a variant *ŋ̍i^{31}* (二). However, these two words have distinct distributions in parallel with Mandarin *er^{51}* (二) and *liang^{214}* (两). Shaoxing Wu '*ŋ̍i^{31}*' as the counterpart of Mandarin '*er^{51}*' is normally used to count, as in (13), or to express an ordinal number, as in (14). Meanwhile, *liaŋ^{113}* is adopted to denote the number two, thus forming so-called numeral classifier phrases of the form [liaŋ^{113}-CL-NP] together with classifiers, as illustrated in (15). In these above situations, *liaŋ^{113}* and ŋ̍i^{31} are not interchangeable.

(13) ie^{ʔ4}, ŋ̍i^{31}/*liaŋ^{113}, sæn^{41}, …
one, two, three,…

(14) di^{113} ŋ̍i^{31}/*liaŋ^{113}
second

(15) liaŋ^{113} (*ŋ̍i^{31}) gə ŋ̍iŋ^{23} /piŋ^{23} ku^{11}
two (*two) CL person/ apple

3.2 Multiple interpretations of liaŋ^{22}-CL-NPs

Researchers have reported that the cardinal number 'two' in many Chinese dialects can also function

[3] The same applies to establishing referents. According to Hawkins (1991), establishing referents in English are normally expressed by a head noun with a following relative clause. Chinese languages, especially in their spoken forms, rarely use these structures. Such usage normally appears in written Mandarin, which is argued to be influenced by European languages, and a reverse word order will be used (Wang 1944 and many others). For example:

Mandarin

(1) Wo^{214} zuo^{35}tian^{55} mai^{214}-de^{0} (na^{51} pen^{35}) hua^{55} si^{214} le^{0}.
I yesterday buy-DE that CL flower dead SFP
'The flower I bought yesterday was dead.'

Shaoxing Wu

(2) ŋo^{113} maŋ^{113} kɒ^{41} dɤ^{231} ma^{113} ge^{0} **(liaŋ^{22} pẽn^{335}) ɕy^{41}**, noʔ^{23} hɒ^{335} kʰø̃^{33}-kʰø̃^{33} goʔ^{0}.
I online buy DE (a few CL) book you may look Prt
'You may take a look the book I bought online.'

as an approximative that means 'a few/a couple of' (Lü, 1999; Tsai, 2002; Wang Xiaohui and Wang Xiaodong, 2012; Sheng, 2019). One instance is its use in Mandarin. In addition to the well-known approximative number *ji*^{214} (some), the cardinal number *liang*^{214} (two) can also be used as an approximative denoting 'a few'. However, such usage of liang-CL-NP in Mandarin is restricted. The approximate interpretation is available only when 1) the tone of 'liang' is neutral and 2) the 'liang-CL-NP' phrase appears postverbally, as in (16). Otherwise, 'liang' can only be interpreted as expressing the exact quantity 'two', as in (17) below.

Mandarin liang (two)

(16) Lai^{35} ke^{51}ren^{0} le^{0} chao^{214} liang^{0} ge^{0} cai^{51} qu^{0}. (neutral tone and postverbal)
come guest FSP fried some dish go
We have guests and I will go to cook some dishes.

(17) Ta^{55} chao^{55}-le^{0} liang^{214}-ge^{51} cai^{51}. (without neutral tone)
she fried-PFV two-CL dish
She cooked two dishes.

(Wang Xiaohui and Wang Xiaodong, 2012)

Differing from Mandarin, the approximative reading of lian^{22}-CL-NP in Shaoxing Wu is available whenever it appears preverbally or postverbally. However, it is normally interpreted as definite or specific in preverbal position, as in (18). Meanwhile, in postverbal position, it is usually understood as indefinite, as in (19):

(18) ga^{11} ɕy^{23}to^{41} ɦo^{11}vɛ^{11} zə^{ʔ45} gə^{ʔ23} **lian^{22}-gəʔ ɕiɒ^{335}niŋ^{23}** na^{ʔ23}ge^{ʔ23} tɕ^{h}ɪ^{ʔ45}-ləʔ kuɑŋ^{41} ga^{11}?
So many dishes very a few-CL kid how eat-PFV finish Prt
How can **these few kids** eat up so many dishes?

(19) piŋ^{23}ku^{11} ŋo^{113} t^{h}a^{ʔ33}gə^{ʔ23} to ^{41} dzɛ^{23}, no^{ʔ23} do^{231} **lian^{22}-gəʔ^{23}** tɕ^{h}i^{33}.
apple I too many Prt, you take a few-CL go
I have too many apples and you may take **some**.

Moreover, 'lian^{22}-CL-NP' is preferred in cases such as anaphoric definite reference, as in (20), bridging cross-reference, as in (21), and reference to salient visible entities, as in (22). 'Lian^{22}-CL-NP' is used for definite reference instead of bare CL-NPs and bare noun phrases.

Anaphoric reference

(20) tso^{33} vəʔ^{23} dʑi^{231} ma^{ʔ0} ŋo^{113} tɕ^{h}iŋ^{335}-ləʔ^{0} lian^{22}gəʔ^{23} ɳiŋ^{231}. pieʔ^{41} ku^{11}, **lian^{22}-gəʔ^{0} ɳiŋ^{231}**
do not finish Prt I hire-PFV a few-CL person. but a few-CL person
(*ɳiŋ^{231}/ *keʔ^{45} lian^{22}-gəʔ^{23} ɳiŋ^{231}) tu^{33} vəʔ^{23} naʔ^{23} geʔ^{23} ɦuɛ^{41} tso^{33} la^{0}.
(*person/*these a few-CL person) all not very can do Prt
'I cannot finish everything myself alone, so I hired **some workers** to help me. However, **the workers** are not helpful'.

Bridging reference

(21) noʔ^{23} ma^{113}ləʔ^{0} bu^{11} ɳi^{31}sɤ^{335} ts^{h}o^{41} a^{0}, **lian^{22}-gəʔ^{23} luẽ^{113}t^{h}ɛ^{41}** kɑŋ^{41}kɑŋ^{41} ɦuø̃^{11}-ku^{11} da^{23}
you buy-Asp CL 2-hand car Prt, a few-CL tires just change-EXP Prt
lɛ^{0}, mæ̃^{11}mæ̃^{11} ɕiŋ^{41}-da^{23} lɛ^{0}
Prt very new-Asp Prt
Did you buy a second car? **The tires** seem to have just been changed and look brand new.

Salient visible reference

(22) ɦo^{231} tsẽ^{41}-tɑŋ^{41} zᴇ^{23}-ku^{41} a^{0}, **liaŋ^{22}-gə^{ʔ0} ɕiɒ^{335}-niŋ^{23}** (***ɕiɒ^{335}-niŋ^{23}** /* **ke^{ʔ45} liaŋ^{22} gə^{ʔ0}**
Prt truly pity Prt a few-CL kids (*kids /*these some-CL

ɕiɒ^{335} niŋ^{23}) ha^{ʔ33}-sa^{ʔ33} dzᴇ^{23} a^{0}!
kids) scare-dead Prt Prt
What a pity! **The children** were shocked by the accident.

Therefore, the interpretation of liaŋ^{22}-CL-NP in Shaoxing Wu seems to be sensitive to the syntactic position where it appears. For instance, it will be understood as indefinite approximative, which is equivalent to 'a few/a couple of' in English, when it occupies the postverbal position. When the liaŋ^{22}-CL-NP appears preverbally as the subject or topic, it will be understood as definite, representing definite, specific plural entities. In the following section, we will examine examples of preverbal liaŋ^{22}-CL-NPs and their alternations with bare nouns and demonstrative phrases. All examples below are from daily conversations by Shaoxing Wu speakers.

4. Plural definite reference and liaŋ^{22}-CL-NPs in Shaoxing Wu

Following Simpson (2017), we will examine the behaviours of definite 'liaŋ^{22}-CL-NP' in four different contexts where plural definite reference is involved: (1) anaphoric context, (2) bridging or associative anaphora context, (3) context with referents (non)visible and with(out) contrast, and (4) context with (non)speaker's kin and personal relations. We will also include the alternations between 'liaŋ^{22}-CL-NP' and other patterns, such as bare nouns, CL-NPs and Dem-CL-NPs, to represent different types of definite reference, which have rarely been mentioned in the literature.

4.1 Anaphoric plural definite reference

One major type of definite reference, as proposed by Hawkins (1978), is the anaphoric definite, where a novel referent is first introduced into a discourse and the following mention must refer back to this referent with a definite noun phrase. The examples below from daily conversions by Shaoxing Wu speakers show that 'liaŋ^{22}-CL-NP' is employed to represent plural anaphoric definite reference.

(23) Context A: A conversation via phone between a mother and her son. The mother complains about her newly hired workers in their family factory.
tso^{33} və^{ʔ23} dʑi^{231} ma^{231ʔ} ŋo^{113} tɕ^{h}iŋ^{335} lə^{ʔ} **liaŋ^{22}-gə^{ʔ23} ɳiŋ^{231}**. piɐʔ^{41} ku^{11}, **liaŋ^{22}-gə^{ʔ} ɳiŋ^{231}**
do not finish Prt I hire-PFV a few-CL person. but a few-CL persons

(***ɳiŋ^{231}** /***ke^{ʔ45} liaŋ^{22} gə^{ʔ23} ɳiŋ^{231}**) tu^{33} və^{ʔ23} na^{ʔ23}-ge^{ʔ23} ɦuᴇ^{41} tso^{33} la.
(*persons/*these some persons) all not very can do Prt
I cannot finish everything by myself, so I hired **some workers** to help me. However, **the workers** are not very good at it.

(24) Context B: A conversation via phone between a mother and her daughter. The daughter sent a parcel to her mother, which had been mentioned to the mother earlier. They are now talking about the presents they have received from each other.
A: m̩^{41} ma^{41}, **gə^{ʔ23}** pɒ^{41}ku^{335} ka^{41} k^{h}ua^{33} tɑŋ^{41} tɒ^{33}dzᴇ^{23} a. noʔ^{23} liaŋ^{22}-dʑiẽ^{11}
mum, CL parcel so quick soon arrive-Asp Prt. you a few-CL

i^{41}zɑŋ^{231} ɦiɤ^{113} ɦiɤ^{113} sɻ^{33}ku^{33}? [4]
clothes have have try-Exp
'Mum, you have received the parcel so quickly. Have you tried the clothes?

B: zɑŋ^{113}ŋ̍a^{11} tɒ^{33} go^{ʔ}. tu^{231} sɻ^{33}ku^{33} dzᴇ^{23}, hɒ^{335} ts^{h}uə̃^{41} go^{ʔ}. **liaŋ^{22}dʑiẽ^{11} i^{41}zɑŋ^{231}**
yesterday arrive Prt all tried-Exp Prt good wear Prt. a few-CL clothes

/***i^{41} zaŋ^{231}*** do^{11}ɕiɒ^{335} ɦa^{113} kɑŋ^{41}kɑŋ^{41} hɒ^{335}
/* clothes size also just good
'I received the parcel yesterday. I have tried all the clothes. **The clothes** are suitable for me in size'.

(25) Context C: *pa^{41}nian^{13}* is a tradition in Shaoxing Wu when relatives and friends visit each other and gather together to celebrate the Chinese Lunar New Year. The gathering party will begin soon. The host is now checking the arrival of guests and finds that all relatives have arrived with the exception of the uncles. A conversion occurred then between the daughter and her mother as the party host is talking about the absence of the uncle, aunt and their son.

A: m̩^{41} ma^{41}, ɕiɒ^{113} dʑiɤ^{11}dʑiɤ^{11}la^{41} væ^{231} veʔ^{2} le^{231}doŋ^{11} le^{231} a^{0}?
mum, little uncle-PL still not come-here come Q
'Mum, have the uncles not arrived here yet?'

B: ŋ̍iɤ^{23}, ŋo^{113} kɑŋ^{41}kɑŋ^{41} ɦia^{231} u^{231}li^{113} tɕy^{11}le^{231}. vəʔ^{2} ɕiɒ^{335}tə^{231} **liaŋ^{22}-gəʔ^{0} ŋ̍iŋ^{231}**
no, I just their house back. not know some-CL persons

/* ***ŋ̍iŋ^{23}*** ta^{33} ɦa^{113}li^{113} tɕ^{h}i^{33} haŋ^{41} dzᴇ^{23}?
/*persons to where go there Prt
'No. I just came back from their house. I did not know where **the persons** were going'.

The phrases *liaŋ^{22}-gəʔ ŋ̍iŋ^{23}* 'two persons' and *liaŋ^{22}-dziẽ^{11} i^{41}zaŋ^{231}* 'two pieces of clothing' in (23)-(25) do not mean 'the two of them'. Rather, the 'liaŋ^{22}-CL-NPs' in all the instances appearing in the preverbal positions denote definite referents that have been mentioned in the earlier discourse: the workers, the clothes and the uncles. In contrast, bare NPs are not allowed in the above situations. As the translations suggest, the most acceptable English correspondence should be 'the NPs', such as 'the workers' in (23), 'the clothes' in (24b) and 'the persons' in (25). Meanwhile, in Mandarin, the demonstrative phrases 'zhexie' (these)-NP/'naxie' (those)-NP will be employed in such contexts.

4.2 Bridging or associative anaphora

The status of bridging or associative anaphora uses of definiteness is arguable in the literature. These cases share features in common with the anaphoric uses and situational/global unique definiteness in some respects (Hawkins, 1978; Schwarz, 2013; Ortmann, 2014; Simpson, 2017 and others). Traditionally, this special type of definite uses is classified into two subtypes of definiteness. (1) The first subtype is the part-whole association, where the definite reference in the second sentence instantiates a part of the preceding referent in the first utterance, as in (26). (2) The second subtype is

[4] Utterance A here demonstrates a bridging context where 'these clothes' are packed in the parcel. Hence, the pattern of liaŋ^{22}-CL-NP denoting plural definite reference is used. We will discuss this type of definite in detail in section 4.2.

relational anaphora bridging, in which the definite has no part-whole relation with the preceding referent but normally has a producer-product relationship, as in (27). The reference in the second sentences in both cases is tied to particular referents that have been mentioned in the preceding context and is thus indirectly anaphoric.

(26) I am going to return my new **laptop** to the store. **The screen** is already broken.
(27) Last year, I bought a beautiful **painting** of Amiens. **The artist** is now quite well known.
(Simpson 2017: 317)

According to Simpson (2017), the definiteness of *the screen* and *the artist* in (26) and (27), respectively, result from an anaphoric-like dependency on the antecedents 'my new laptop' and 'a beautiful painting', respectively. Such definites relate back to the context in an indirect anaphoric way. The definite does not refer back to any painting or laptop but is rather understood to refer to the unique painting or laptop discussed in the preceding utterances. Therefore, both anaphoric and unique strategies exist in the bridging or associative anaphora uses of definite descriptions. Cross-linguistically, different determiners are employed to represent these two cases of bridging or associative anaphora definiteness. For instance, in German, a weak determiner is used for relational bridging in the same way as for typical anaphoric definiteness; meanwhile, a strong determiner is used in instances of producer-product bridging, thus resembling the representation given to situational/global unique referents (Schwarz, 2009).

Simpson (2017) reported that definiteness marking in Jinyun Wu does not distinguish these so-called two subtypes of bridging, namely, part-whole and producer-product, but predominantly uses the bare classifier strategy in bridging situations. Intuitively, Shaoxing Wu distinguishes producer-product bridging from the part-whole relationship by adopting different definiteness strategies. For the cases of producer-product, the bare classifier structure is used when a singular reference is mentioned, and 'lian^{22}-Cl-NP' will be employed when the reference is plural, as in (28)-(29) below. For the cases of part-whole bridging, both CL-NPs and bare NPs are allowed. However, the more specific the reference, the more the classifier strategy will be preferred, as in (30)-(31) below.

Producer-product relation

(28) A: ɦia^{113} zɑŋ^{113}ŋa^{11} tɕʰi^{33} kʰæn^{33}lə^{ʔ} **go^{11} tsæn^{335} læ̃^{113}**, kʰæn^{33} ɦuo^{33}ɦuo^{33}.
we yesterday go watch-PFV CL exhibition, watch painting.
'Yesterday, we went to a painting exhibition'.

B: hɒ^{335}hɒ kʰæn^{33} ?
good watch?
'Were the paintings good?'

A: **<u>sə^{ʔ45} ɦuo</u>** dɔ^{11} du^{231} mæ̃ ^{335}vəʔ^{2}tsʰo^{33}, pie^{ʔ45} ku^{11} **<u>lian^{22} gə^{ʔ} ɦuo^{33}tɕia^{41}/ * ɦua^{33} tɕia^{41}</u>**
CL(PL) painting but all very not bad, but some-CL painter /*painter

tu^{41} ve^{ʔ2} na^{ʔ23}ge^{ʔ23} tʰiŋ^{41}ŋiẽ^{231}-ku go^{ʔ}
all not how hear-EXP Prt
'All **<u>the pictures</u>** were good, but we had never heard of **<u>the painters</u>** before'.

(29) A: ŋo^{113} mɑŋ^{113} kɒ^{41} dɤ^{231} ma^{113} lə^{ʔ} **<u>lian^{22} pẽn^{335} ɕy^{41}</u>** 。
I online buy-Asp some CL book
'I bought **<u>a few books</u>** online'.

B：hɒ^{335}hɒ kʰæn^{33} ?
good looking?
'Are they interesting?'

A: væ^{231} hɒ^{335}. pieʔ^{45} ku^{11} **<u>lianŋ^{22}gəʔ tsoʔ^{33}tsɛ^{335} /*tsoʔ^{33}tsɛ^{335}</u>** kʰo^{335}nɛ̃^{231} noʔ^{23}
still good. However, a few-CL author /*author maybe you

veʔ^{22} naʔ^{23}geʔ^{23} tʰiŋ^{41}ɳiɛ̃^{231}-ku goʔ.
not how hear-EXP SFP
'Not bad. However, you might have never heard of **<u>the authors</u>**'.

In (28), the term means 'the relevant artists' and not 'two artists'. Similarly, in (29), a product-producer relationship holds between the referent of 'lianŋ^{22}-CL-NP' and its antecedent paintings because 'the painters' are understood as the particular painters who painted the paintings on display in the exhibition that the speakers are discussing. In all these cases, the bare NP strategy is ruled out in Shaoxing Wu.

Part-whole relation

(30) noʔ^{23} ma^{113}-ləʔ **<u>bu^{11} ɳi^{23} sɤ^{335} tsʰo^{41}</u>** a, **<u>lianŋ^{22}-gəʔ luɛ̃^{231} tʰɛ^{41}/ luɛ̃^{231} tʰɛ^{41}</u>**
you buy-PFV CL second-hand car Prt, a few-CL tires / tires

kɑŋ^{41}-kɑŋ^{41} ɦuø̃^{11}- ku da^{41}-lɛ, mæ̃^{231}-mæ̃^{231} ɕiŋ^{41} da^{41}-lɛ
just change-EXP SPF, very-very new SFP
'Did you buy **<u>a second-hand car</u>**? **<u>The tires</u>** seem to have just been changed and look brand new'.

(31) noʔ^{23} ma^{113}-ləʔ bu^{11} ɳi^{23}-sɤ^{335} tsʰo^{41} a, **<u>bu^{11} faʔ^{41}-doŋ^{11}-tɕi^{41}/?faʔ^{41}-doŋ^{11}-tɕi^{41}</u>**
you buy-PFV CL second-hand car Prt CL engine /??engine
kɑŋ^{41}- kɑŋ^{41} ɦuø̃^{11}-ku da^{41}-lɛ, mæ̃^{231}-mæ̃^{231} ɕiŋ^{41} da^{41}-lɛ
just change-EXP SFP, very-very new SFP
'Did you buy **<u>a second-hand car</u>**? **<u>The engine</u>** seems to have just been changed and looks brand new'.

Similarly, the utterance in (30) denotes a part-whole relationship in which the tires are naturally understood as the tires of the second hand car that the speakers are discussing. Given that a car normally has more than one tire, both the 'lianŋ^{22}-CL-NP' and bare NP are acceptable to denote the plural reference. However, differences in meaning between these two options exist. For the option of 'lianŋ^{22}-CL-NP', the speaker emphasises each tire, or at least some of them share the property of being new. In contrast, a bare NP will be adopted if the speaker aims to comment on the tires in general. When we change the tires to the engine, as in (31), which is a singular referent, the bare classifier strategy is preferred because each car has only one engine. Such an alternation related to the specificity of the reference can be illustrated further in cases such as (32) below.

(32) a: noʔ^{23} ɦuo^{11} **<u>dʑiɛ̃^{11}ʔi^{41}zɑŋ^{231}</u>** hɒ^{335}-vəʔ^{23}-hɒ^{335} ?
you say CL cloth good-not-good?
'Do you think **<u>the clothing</u>** is good?'b: hɒ^{335} goʔ. **<u>tsəʔ^{45} ɳæ̃^{235}səʔ^{33} /ɳæ̃^{235}səʔ^{33}</u>** dəʔ^{11}bɪʔ^{231} hɒ^{235}. pieʔ^{41}ku^{11}
good Prt. CL colour /colour especially good. but

<u>lian^{22} tsə^{ʔ45} ɕiɤ^{33} kʰɤ^{335}</u> / **<u>ɕiɤ^{33}kʰɤ^{335}</u>** hɒ^{335}ɦian^{11} dzaŋ^{23}-lə^{ʔ} sə^{ʔ45}
a few-CL sleeve / sleeve seem long-PFV a bit.
'It's good, especially **<u>the colour</u>**. **<u>The sleeves</u>** seem to be a bit long'.

b': **<u>tsə^{ʔ45} ŋæ̃^{235}sə^{ʔ33} /*ŋæ̃^{235}sə^{ʔ33}</u>** mæ̃^{231} pʰiɒ^{33}lian^{11}, ŋo^{113} də^{ʔ11}bɪ^{ʔ231} huə̃^{41}ɕi^{335}.
CL colour /*colour very beautiful I especially like

pie^{ʔ41} ku^{11}, ɦi^{231} **<u>lian^{22}-tsə^{ʔ45} ɕiɤ^{33}kʰɤ^{335} /*ɕiɤkʰɤ</u>** hɒ^{335}ɦian^{11} dzaŋ^{23}-lə^{ʔ} sə^{ʔ45}.
but its a few-CL sleeve /*sleeve seem long-PFV a bit
'I like **<u>the colour</u>** the most, but its **<u>sleeves</u>** seem to be a bit long'.

In (32b), when bare NPs are used, it is understood that the piece of clothing has more than one colour and the sleeves in general are rather long. No such reading is available when classifier-NPs are used. In contrast, when the clothing clearly has only one colour and the speaker wants to emphasise the length of the clothing's sleeves by using an appositive structure, only the classifier-NP is acceptable, as in (32b'). In addition to the appositive structure where a bare NP strategy will always be ruled out, cases with modifiers also only allow the CL-NP structure, as exemplified in (35)-(36) below. The contrast in acceptability between (31)-(32) and (33)-(34) shows that the use of the CL-NP pattern for singular definite reference may override the use of bare NP due to the specificity constraint. In (33)-(34), the degree of specificity is strengthened by adding modifiers to narrow down the range of reference.

(33) zə^{ʔ11} ka^{41} ɕiŋ^{41} **<u>bu^{11} fa^{ʔ41}-doŋ^{11}-tɕi^{41} /*fa^{ʔ41}-doŋ^{11}-tɕi^{41}</u>** a^{0}, kɑŋ^{41}-kɑŋ^{41} ɦuə̃^{11}-ku^{0} go^{ʔ23}.
truly very new CL engine / *engine Prt just change-EXP Prt
'**<u>The engine</u>** is so new. It must have just been changed'.

(34) zə^{ʔ11} ka^{41} ɕiŋ^{41} **lian^{22}-gə^{ʔ} <u>luẽ^{231} tʰɛ^{41}/* luẽ^{231} tʰɛ^{41}</u> a^{0}**, kɑŋ^{41}-kɑŋ^{41} ɦuə̃^{11}-ku^{0} go^{ʔ23}.
truly very new a few-CL tires / * tires Prt just change-EXP Prt
'**<u>The tires</u>** are so new. they must have just been changed'.

4.3 Referents (non)visible and with(out) contrast

According to Simpson (2017), the visibility of the referent also plays a significant role in determining the use of bare classifier and bare noun patterns in Jinyun Wu. If the referent is present in front of the speaker and listener, this will naturally result in the use of a bare classifier form, as in (35a), where the classifier performs a deictic, demonstrative-like function. When the speaker emphasises the absence of a particular referent, this will regularly result in a bare noun pattern, as in (35b).

Jinyun Wu

(35) a. Context: The speaker and his family members hear a dog barking outside their house and wonder if it is their (unique) dog. The speaker says:
<u>gɔdʑiəm/#dzaɪ gɔdʑiəm</u> leɪ tɕieɪkeɪ a?
dog /#CL dog is where Prt
'Where is **<u>the dog</u>**?'

b. Context: The speaker and his family are looking for their missing dog. When they finally see it, the speaker says, 'Look! The dog is on the roof of the house'.
niɔ a! **<u>dzaɪ gɔdʑiəm /#gɔdʑiəm</u>** leɪ au-dan dʑiɔ.

look Prt CL dog /#dog is roof top
'Look! **The dog** is on the roof of the house'.

However, these predictions are not borne out in Shaoxing Wu, as shown in (36) below, where the equivalent of (35) in Shaoxing Wu demonstrates that the bare classifier is always the natural choice even when the referent is absent from the scene, as long as specific entities are involved.

Shaoxing Wu

(36) a. Context: The speaker and his family members hear a dog barking outside their house and wonder if it is their (unique) dog. The speaker says:

tsəʔ^{23} kɤ^{335}/* kɤ^{335} ta^{33} ɦa^{231}-li^{113} tɕʰi^{33} dzɛ^{231}?
CL dog /*dog go where go Prt
'Where is **the dog**?'

b. Context: The speaker and his family are looking for their missing dog. When they finally see it, the speaker says, 'Look! The dog is on the roof of the house'.

nɒ^{113}, **tsəʔ^{23} kɤ^{335}/* kɤ^{335}** ʔuoʔ^{23}tiŋ^{41} kɒ^{41}-dɤ^{231} lɛ^{231} da^{0}.
Prt, CL dog /*dog roof upside here Prt
'Look! **The dog** is on the roof of the house'.

The above data in Shaoxing Wu demonstrate that whether the referent is present or absent, the bare classifier form *'tsəʔ kɤ'* (the dog) is preferred as long as the specific single individual/entity being discussed is known to the speaker and/or listener, in the sense of Himmelmann's (1996) 'recognitional referents' or Hawkins's (1978) immediate-situation definites. Therefore, factors other than the visibility of referents influence the choice of bare classifier and bare noun forms to represent singular definite reference in Shaoxing Wu. Moreover, what if the reference is plural or mass? Will the presence/absence of the referents also influence the choice of liaŋ^{22}-CL-NP and bare noun patterns in Shaoxing Wu? Simpson (2017) did not tackle these questions. Thus, we consider the following examples first.

Shaoxing Wu

(37) Context: A car has crashed into a wall. The speaker and a friend are walking by, and they hear that kids who are crying hard. They stop to look at the car, and the speaker says:

ɦo^{11} tsẽ^{41}-tɑŋ^{41} zɛ^{231}-ku^{41} a^{0}, **liaŋ^{22}-gəʔ^{0} ɕiɒ^{335}-niŋ^{23} (*ge^{33} liaŋ^{22}-gəʔ^{0} ɕiɒ^{335}-niŋ^{23}**
Prt truly pity Prt, a few-CL kids (*those a few kids

/*ɕiɒ^{335}- niŋ^{23}) haʔ^{335}-saʔ^{335} dzɛ^{23}.
/ *kids) scare- dead Prt
'What a pity! **The kids** were shocked by the accident'.

(38) Context: A car has crashed into a wall, injuring everyone in the car. All patients were sent to the hospital, leaving a large amount of blood on the street. The speaker and a friend are walking by, and they see kids who are shocked and have fainted. They stop, and the speaker says:

ɦo^{11} tsẽ^{41}-tɑŋ^{41} zɛ^{231}-ku^{41} a^{0}, **ɕiɒ^{335}-niŋ^{23} /? liaŋ^{22}gəʔ^{0} ɕiɒ^{335}-niŋ^{23}**
Prt truly pity Prt kids / ? a few-CL kids

/*ge^{33} liaŋ^{22} gəʔ ɕiɒ^{335}-niŋ^{23} zɛ^{23} haʔ^{33}-saʔ^{335} dzɛ^{23} goʔ^{23}.

/*Dem a few-CL kids surely scare-dead Prt Prt
'What a pity! **The kids** will surely be shocked by the accident'.

In (37), both the speaker and listener are at the site and see that more than one child is shocked by the accident and crying. The speaker draws the listener's attention to the crying reaction of those specific children on the scene. Therefore, 'lian^{22}-CL-NP' is adopted. In contrast, in (38), the speaker draws the listener's attention to the general reaction of a child facing a terrible accident without the emphasis on particular individuals. Thus, a bare noun form is preferred. Since the speaker has no intention to compare the referents with other kids in (37) and (38), the demonstratives kəʔ^{45}/haŋ^{231} (these/those) are not allowed in either instance above. However, examples such as (39) below show that the demonstrative phrase will override the 'lian^{22}-CL-NP' and bare noun forms if the definite referents are highlighted and contrasted.

(39) A: no^{ʔ23} ka^{41} kʰoŋ^{33} a^{0}, pʰo^{41}-lɒ^{113} **lian^{22}-sø̃^{33} tsʰɑŋ^{41}mẽ^{113}** væn^{23}-hɒ^{335}
you very free Prt why a few-CL window why not

kʰa^{41}-kʰa^{41} ɦi^{231} go^{ʔ33}.
wipe it Prt
You are so free. Why don't you wipe **the windows**?

B: ɦia^{231}-li^{113} **lian^{22}-sø̃^{44} tsʰɑŋ^{41}mẽ^{113}** lɒ^{231}, və^{ʔ23}-zə^{ʔ23} mæ̃^{231}-mæ̃^{231} tɕʰiŋ^{41}-sɑŋ^{335} doŋ^{11}?
Which a few-CL window Prt not-is very clean STAT
Which **windows** do you mean are dirty? They are very clean, aren't they?

A：**kəʔ^{45}/haŋ^{231} lian^{22} sø̃^{41} tsʰɑŋ^{41} mẽ^{231}**/* **sø̃ tsʰɑŋ^{41}mẽ^{231}**/* **tsʰɑŋ^{41}mẽ^{113}** lo^{0}.
these/those some windows /* CL windows /* window Prt
These/those windows are dirty.

Differing from Mandarin, the use of Dem-lian-CL-NP in Shaoxing Wu is highly limited and only used where the definite referents are listed in contrast and deictic to the speaker. The unacceptability in examples (37)-(39) also illustrates this limitation.

4.4 (Non)speaker's kin and personal relations

Simpson (2017) argued that the familiarity effect plays an important role in the selection of definite options in Jinyun Wu. Specifically, if speakers have an existing mental representation for an entity and are familiar with it, a bare noun form is frequently used for definite reference. Meanwhile, if the speaker has a nonexistent or very weak mental representation for those objects with which he/she has never come into contact, a bare classifier pattern is preferred. This so-called familiarity effect occurs with reference to family members in Shaoxing Wu. For example, when a mother is talking about her son and daughter, the bare noun form is the most appropriate for reference, as in (40). In contrast, the bare classifier pattern is more natural if she is talking about other kids with whom she is not familiar or has no close relation, as in (41).

(40) ŋa^{113} lian^{113}gəʔ^{3} ɕiɒ^{335}-ɳiŋ^{23}, ʔɪʔ^{45}gəʔ ɳi^{231}tse^{335}, ʔɪʔ^{45} gəʔ^{45} nø̃^{113}. **ɳi^{231} tseʔ^{335}**
we 2-CL kids, 1-CL son, 1-CL daughter. son

/* **gəʔ^{45}ɳi^{231} zeʔ^{335}** ŋ̩^{113} sᴇ^{33} ləʔ^{0} **nø̃^{113}** **/*gəʔ nø̃^{113}** sæ̃^{41} sᴇ^{33}.

/* CL son five years and daughter /*CL daughter three years
'I have two kids: a son and a daughter. **The son** is five years old, and **the daughter** is three years old'.

(41) zɑŋ^{113}ŋa^{11} diẽ^{33}-zɿ^{11} ɕiŋ^{41}-vẽ^{231} kɒ^{41}-dɤ^{231} gə^{ʔ45} ŋiŋ^{23}-ko^{41} tsẽ^{41}-tɑŋ^{41} zɛ^{231}-ku^{41} dzɛ^{23}
yesterday TV news upside CL family really pitiful Prt

ɦa^{113} zə^{ʔ23} ʔɪʔ^{45} gə^{ʔ45} ŋi^{231}-tseʔ^{335} lə^{ʔ0} ʔɪʔ^{45}-gə^{ʔ45} nə̃^{113}. **gə^{ʔ45} ŋi^{231}-tseʔ^{335}**
also is 1-CL son and 1-CL daughter. CL son

tɕ^{h}ɪ^{ʔ45} sɛ^{33} lə^{0} **gəʔ nə̃^{113}** ŋ^{113} sɛ^{33}
seven years and CL daughter five years.

'The family reported by the TV news yesterday was really poor. They also have a son and a daughter. **The son** is seven years old, and **the daughter** is five years old'.

Such distinctions in encoding depending on the degree of familiarity that the speaker has with her family members and others' family members also exist when the definite referents are plural. Specifically, a bare noun will be used when the speaker refers to his/her own family members, while the 'liaŋ^{22}-CL-NP' pattern will be used for definite reference to other people's kin. In (42), for instance, the speaker is talking about her seven kids in (42a) and her neighbour's seven kids in (42b).

(42) a. ŋa^{113} tɕ^{h}ɪ^{ʔ45} gə^{ʔ45} ɕiɒ^{335}-ŋiŋ^{23}, sæ̃^{41} gə^{ʔ45} ŋi^{231}-tseʔ^{335} sɿ^{33} gə^{ʔ45} nə̃^{113}.
we seven CL kids, three CL son four CL daughter

ŋi^{231}-zeʔ^{335} məʔ^{0}/* **liaŋ^{22}-gəʔ^{0} ŋi^{231}-tseʔ^{335}** məʔ^{0} tu^{41} zəʔ^{23} ʔi^{231}-saŋ^{41}, **nə̃^{113}** məʔ^{0}
son Prt /* a few CL son Prt all are doctor, daughter Prt

/***liaŋ^{33}-gəʔ^{3} nə̃^{113}** məʔ^{0} tu^{41} zəʔ^{23} lɒ^{113}sɿ^{41}.
/* a few-CL daughter Prt all are teacher

'I have seven kids, three sons and four daughters. **The sons** are all doctors, and **the daughters** are all teachers'.

b. ɦia^{113} u^{231}-li^{113} tɕ^{h}ɪ^{ʔ45} gə^{ʔ41} ɕiɒ^{335}-ŋiŋ^{23}, sæ̃^{41} gə^{ʔ45} ŋi^{231}-tseʔ^{335} sɿ^{33} gə^{ʔ45} nə̃^{113}.
they family seven CL kids, three CL son four CL daughter.

liaŋ^{22}-gəʔ ŋi^{231}-tseʔ^{335} /*ŋi^{231}-tseʔ^{335} məʔ^{0} tu^{41} zəʔ^{23} ʔi^{231}-saŋ^{41}, **liaŋ^{33}-gəʔ^{3} nə̃^{113}**
a few-CL son / *son Prt all are doctor a few-CL daughter
/* nə̃^{113} məʔ^{0} tu^{41} zəʔ^{23} lɒ^{113}sɿ^{41}.
/*daughter Prt all are teacher

'They have seven kids, three sons and four daughters. **The sons** are all doctors, and **the daughters** are all teachers'.

Next we consider inanimate reference. According to Simpson (2017), the so-called familiarity effect will override other factors in determining the definite uses in Jinyun Wu, whether the reference is animate or inanimate. As mentioned above, however, Shaoxing Wu does not behave in the same way as Jinyun Wu when the referents are inanimate. For example:

(43) A: ŋa^{113} zɑŋ^{113}ŋa^{11} tɕʰi^{33} ma^{113}-lə^{ˀ0} sə^{ˀ45} diẽ^{11}tɕʰi^{41}, ʔɪʔ^{23}-bu^{33} diẽ^{11}zɻ^{23}tɕi^{41}
we yesterday go buy-PFV some electricity one-CL TV set

lə^{ʔ} ʔɪʔ^{23}-bu^{33} diẽ^{11} nɒ^{335}.
and one-CL computer.
We bought a TV set and a computer yesterday.

B: no^{ˀ23} **bu^{33} diẽ^{11}zɻ^{23}tɕi^{41}** ʔiɒ^{23} ɦuə̃^{11}-ku^{33} dzᴇ^{23} ɦa^{23}?
you CL TV set need replace-Exp Prt Q?
Did **the TV set** need to be replaced?

A: **bu^{33} dziɤ^{11} diẽ^{11}zɻ^{23}tɕi^{41}** pʰa^{33}-ɦuæ̃^{11} dzᴇ^{23} a^{0}, ma^{113}-lə^{ˀ0} bu^{33} ɕiŋ^{41} go^{ˀ0}.
CL old TV set broken Prt Prt, buy-PFV CL new Prt.
bu^{33} diẽ^{11}-nɒ^{335} mə^{ʔ0} zɑŋ^{11}-kʰo^{33} ʔiɒ^{23} ɦioŋ^{11} go^{ʔ0}
CL computer Prt take-class need use Prt
'**The old TV set** is out of order, so we bought a new one. The computer is used for taking online classes.'

(44) A：ɦia^{113} tɕʰi^{33} ma^{113} diẽ^{11}zɻ^{23}tɕi^{41} tɕʰi^{41}-haŋ^{231}
they went buy TV set go there
'They went to buy a TV set'.

B: pʰo^{41}-lɒ^{113} **bu^{33} diẽ^{11}zɻ^{23}tɕi^{41}** pʰa^{33}-ɦuæ̃^{11} haŋ^{41} dzᴇ^{23} a^{0}?
maybe CL TV set broken there Prt Q
'Was **the TV set** out of order?'

A: və^{ˀ23} ɕiɒ^{335}-tə^{ˀ45}, ʔɒ^{33}mɒ^{11} zᴇ^{11}. ɦi^{231} ɦuo^{11} **bu^{33} dziɤ^{11} diẽ^{11}zɻ^{23}tɕi^{41}** fɑŋ^{33} la^{231}
not know maybe is she said CL old TV set put in

fɑŋ^{335}kæ̃^{41} li^{113}, **ɕiŋ^{41} diẽ^{11}zɻ^{23}tɕi^{41}** ma^{113}-lᴇ^{0} hɒ^{335} fɑŋ^{33} la^{11} kʰa^{ʔ23} tʰiŋ^{41} li^{335}.
bedroom inside new TV set buy-PFV can put in sitting-room inside
'I have no idea. It may be yes. She said **the old TV set** can be moved to the bedroom and **the new one** will be placed in the sitting room'.

In (43), the speaker has a clear mental representation of 'the TV set and the computer' because she participated in purchasing these objects. However, the bare classifier pattern of '*bu^{33} dziɤ^{11} diẽ^{11} zɻ^{23} tɕi^{41}*' (CL-old TV set) and '*bu^{33} diẽ^{11} nɒ^{335}*'(CL-computer) is used in the last utterance, as in other typical anaphoric contexts shown in (23)-(25). In contrast, the speaker and the listener may or may not have a mental representation of their friend's old TV set. However, the CL-NP is used. Meanwhile, a bare NP 'new TV set' is adopted to encode a referent that both discourse participants have not previously seen. As in the other contexts mentioned above, when an inanimate referent is involved, Shaoxing Wu prefers the classifier-NP structure as long as the reference is specific or the speaker wants to emphasise the specificity of the referent they are discussing, which is different from the case of animate reference. Such an effect also exists in the case of plural definite reference, as in (45)-(46) below.

(45) ŋo^{231} mɑŋ^{113}-kɒ^{41}-dɤ^{231} ma^{113}-lə^{ˀ0} liaŋ^{22}-pẽ^{41} ɕy^{41}, ŋo^{231} tu^{41} kʰø̃^{33}ku^{41} zᴇ^{23},
I online buy-PFV a few-CL book I all see-Exp Prt

vəʔ^{223} tsʰo^{33} goʔ^{0}. no^{ʔ23} hɒ^{335} kʰø̃^{33}-kʰø̃^{33} goʔ^{0}. **<u>liaŋ^{22} pẽ^{41} ɕy^{41}</u>** fɑŋ^{231}kæ̃^{41} ɕy^{41}ko^{33}
not bad Prt you may see Prt a few-CL book bedroom shelf

kɒ^{41}-dɤ^{231} lɛ^{0} doŋ^{11}.
upside at there
'I bought some books online, and I have read them. The books are good. You may take a look. **<u>The books</u>** are on the shelf in the bedroom'.

(46) ɦi^{231} ɦuo^{11} ɦi^{231} mɑŋ^{113}-kɒ^{41}-dɤ^{231} ma^{113}-lə^{ʔ0} liaŋ^{22}-pẽ^{41} ɕy^{41}, ɦi^{231} ko^{ʔ231}-tə^{ʔ0}
she say she online buy-PFV a few-CL book she think

vəʔ^{223} tsʰo^{33}. ɦi^{231} ɦuo^{11} **<u>liaŋ^{22}-pẽ^{41} ɕy^{41}</u>** fɑŋ^{231}kæ̃^{41} ɕy^{41}ko^{33} kɒ^{41}-dɤ^{231} lɛ^{0} doŋ^{11}.
not bad she say **a few-CL book** bedroom shelf upside at there
no^{ʔ23} hɒ^{335} kʰø̃^{33}-kʰø̃^{33} goʔ^{0}.
you may see Prt
'She said she bought some books online and she thought they are good. She said **<u>the books</u>** are on the shelf in the bedroom and you may also take a look'.

Similar to the instances involving singular reference, '*liaŋ^{22}pẽ ɕy*' (2-CL book) is employed in (45) and (46) regardless of the 'familiarity' toward the books. Therefore, the so-called familiarity effect exists only in cases of animate reference to those individuals with whom the speaker is most familiar, such as his/her own family members, particularly in Shaoxing Wu. Unlike Jinyun Wu, the degree of 'specificity' of the reference plays a large role in the selection of bare NP versus classifier patterns in cases of inanimate reference.

4.5 Interim conclusions

Our discussion illustrates that Shaoxing Wu uses the 'liaŋ^{22}-CL-NP' string to represent plural definite reference. It is predominantly employed in cases of anaphoric definite, producer-product bridging reference and salient reference without contrast and animate reference involving the speaker's personal relations. Moreover, it alternates with the bare NP strategy in situations of part-whole bridging reference and inanimate reference with or without the speaker's personal relations. The more specific the definite reference is, the more 'liaŋ^{22}-CL-NP' will be preferred. However, 'liaŋ^{22}-CL-NP' is interchangeable with the bare NP, especially in part-whole bridging reference. The definite strategies to encode plural reference in Shaoxing Wu can be summarised in Table 1 below.

Table 1: Choice of Liaŋ^{22}-CL-NP and bare NP to represent plural definite reference

Situations where plural definite reference is involved		Definite strategies	
		Liaŋ^{22}-CL-NP	**Bare NP**
Anaphoric definite reference		✓	
Bridging reference and definite associate anaphora	producer-product	✓	
	part-whole	✓ (specific individuals)	✓ (in general)
Salient referent without contrast		✓	

Animate reference	speaker's kin and personal relations		✓
	nonspeaker's kin and without personal relations	✓	
Inanimate reference		✓	

5. 'Liaŋ-CL': A grammaticalisation of (in)definite plural marking strategy

In addition to serving as a cardinal number denoting quantity 'two', as mentioned in section 2, *liaŋ* with a level tone (22 pitch on a 1-5 scale, 1 being the lowest pitch) together with a classifier functions as an approximation ('a few, a couple of') in most cases discussed above. This usage appears to be an instance of a grammaticalisation process where *liaŋ* undergoes semantic bleaching and is grammaticalised into an approximate that denotes a definite or indefinite set with plural members. Such a grammaticalisation pathway from the numeral two to a plural definite marker has not been described extensively in the literature.

5.1 *Liaŋ* 'two' + classifier > DUAL

According to Long et al. (2012) and Kuteva et al. (2019), Mandarin Chinese *liang*^{214} 'two' undergoes grammaticalisation and follows a chain of *liang* 'two'+ *ge*, classifier > *liang* 'two' > *lia*, where '*lia*'(俩) is the combination of '*liang'* and classifier *'ge'* and shows phonological reduction to function as a dual marker for personal pronouns, as shown in (47).

(47) ta^{55}men^{0} lia^{214}/ wo^{214}men^{0} lia^{214}
they DU / we DU
'they two' / 'we two'

We basically agree with the above observation but would add that '*lia*', as a dual marker in Mandarin, may follow not only a pronoun but also other common human nouns, as exemplified in (48), where the common noun jie^{214}mei^{51} (sister) plus lia^{214} means 'two sisters'.

(48) jie^{214}mei^{51} lia^{214} /xiong^{55}di^{51} lia^{214} /fu^{55}qi^{55} lia^{214}
sisters DU /brothers DU /couple DU
'the two sisters/brothers/the couple'

Mandarin '*lia*^{214}' has no exact counterpart of in Shaoxing Wu. To express cases such as in (47)-(48), two distinctive patterns are employed in Shaoxing Wu: one is the 'pronoun (plural)-liaŋ-CL' sequence[5], as in (49) below, where the plural pronoun is followed by the '*liaŋ*^{113}*-CL*', meaning 'pronoun-DU'. The other is the structure of '*liaŋ*^{113}-common human noun', where '*liaŋ*^{113}' does not have a classifier, and is combined directly with a common human noun to denote 'two of X', as illustrated in (50).

(49) ɦia^{22} liaŋ^{113}-gəʔ^{23} /ŋa^{113} liaŋ^{113}-gəʔ^{23}

[5] The sequence of 'pronoun (plural)-liaŋ-CL' in Shaoxing Wu can also mean 'pronoun- two or more' if the pronoun is followed by the approximate 'liaŋ^{22}-CL', where the tone of 'liaŋ' changes into a level one and the numeral meaning is bleached. In addition, the classifier with phonological reduction to the neutral tone attaches to the 'liaŋ', thus forming a tonal unit.

(3) ɦia^{22} liaŋ^{22}-gəʔ^{0} /ŋa^{113} liaŋ^{22}-gəʔ^{0}
they two-CL / we two-CL
'they (two or more)'/'we (two or more)'

they DU / we DU
'they two'/'we two'

(50) liaŋ^{113} tɕi^{335}mɛ^{11}/liaŋ^{113} ɕioŋ^{41}di^{11}/liaŋ^{113} fu^{41}tɕʰi^{41}
DU sisters / DU brothers / DU couple
'the two sisters/brothers/the couple'

Interestingly, the tone of *liaŋ* in all DU cases is rising, as it is in the cardinal number *liaŋ* (two). The classifier cannot be omitted in the pattern of 'pronoun (plural)-liaŋ-CL'. Otherwise, the expressions will be unacceptable, as exemplified in (51). In contrast, the omission of a classifier in cases such as (50) above is obligatory. It will lead to a distinct interpretation if a classifier is inserted, such as *tɛ^{33}* 'pair', as illustrated in (52). Here, '*liaŋ^{113}-tɛ^{33} tɕi^{335}mɛ^{11}*' does not mean 'the two sisters' but 'two pairs of sisters', where four persons in total are involved. This usage is in contrast with cases such as (50) above in which exactly two persons are being referred to.

(51) * ɦia^{22} liaŋ^{113}/* ŋa^{113} liaŋ^{113}
* they DU /* we DU
'they two'/'we two'

(52) liaŋ^{113}-tɛ^{33} tɕi^{335}mɛ^{11}/liaŋ^{113} tɛ^{33} ɕioŋ^{41}di^{11}/liaŋ^{113} tɛ^{33} fu^{41}tɕʰi^{41}
two-CL sisters / two-CL brothers /two-CL couple
'two pairs of sisters/brothers/two couples'

Similar to Mandarin, Shaoxing Wu *liaŋ^{113}* 'two' also undergoes grammaticalisation and follows a chain of liaŋ 'two' + (classifier) > DUAL. The cardinal number 'two' in both of these Sinitic languages has developed a new function of DUAL. Moreover, they only differ from each other in the degree of grammaticalisation, specifically, whether a unified form exists to express DUAL. However, such grammaticalisation of cardinal numbers embraces other languages in the world, such as old English, West !Xun and others. Therefore, Mandarin lia^{214} 'DUAL' and Shaoxing Wu *liaŋ^{113}* 'DUAL' are instances of a more general process where lower numerals are pressed into service to function as number markers (Heine and Kuteva, 2002).

5.2 DUAL > (in)Def PLURAL (APPROXIMATIVE)

As discussed in the above sections, the numeral 'two' in Shaoxing Wu together with a classifier forms the sequence '*liaŋ-CL*', where the lexical meaning of quantity 'two' is lost, and the sequence with an approximate meaning 'a few', where the noun phrase represents (in)definite plural reference. This usage suggests that *liaŋ* 'two' in Shaoxing Wu, which means 'a few/a couple', can also be the source of a plural marker.

Unlike Mandarin, which has its writing system and abundant written resources to witness historical changes, Shaoxing Wu, as a minor dialect in Chinese, has no written form of its own to record the language. Based on the approximative usage of 'liaŋ-CL-NP' in modern Shaoxing Wu, three possible grammaticalisation pathways for liaŋ 'two' can be hypothesized: (1) liaŋ 'two' + classifier > DUAL, (2) liaŋ 'two' + classifier > (in)Def PLURAL (APPROXIMATE), and (3) liaŋ 'two' + classifier > DUAL > (in)Def PLURAL (APPROXIMATIVE). We support the third path, considering all of the abovementioned facts of 'liaŋ-CL' in Shaoxing Wu as well as the following evidence from other

languages. One piece of evidence is from US English, where measure nouns, such as 'a couple of', which denotes 'DUAL', can also be used as an approximative with the meaning of 'a few'. For example:

(53) I had to wait **a couple (of) hours**.

(54) I played with **a couple (of) friends** yesterday.

The phrases 'a couple hours' and 'a couple friends' in sentences (53) and (54), with optional reduction in form from 'a couple of' to 'a couple', means 'two hours/friends or slightly more'. The meaning of 'two' persists, but with an approximative sense. Thus, the measure noun 'a couple (of)' in English undergoes a similar grammaticalisation following a chain of A COUPLE OF 'dual' > APPROXIMATIVE 'a few'.

Another piece of evidence is from the grammaticalisation path of 'three' in the Khoisan language !Xun, as stated in Heine and Kuteva (2002) and Kuteva et al. (2019). Very similarly, a new function PLURAL of the number 'three' stems from TRIAL rather than THREE directly. According to Heine and Kuteva (2002) and Kuteva et al. (2019), the numeral THREE in !Xun evolves into a trial and then a plural marker on personal pronouns and other word categories, thus following the overall chain of grammaticalisation: THREE > TRIAL > PLURAL.

Regarding these two grammaticalisation paths in natural languages, we assume that the APPROXIMATIVE 'PLURAL' in Shaoxing Wu also stems from the DUAL 'liaŋ + (CL),' which evolved from the *liaŋ* 'two' + classifier. Shaoxing Wu would then be another instance of this general grammatical process. Therefore, instead of the grammaticalisation process of 'liaŋ "two" + classifier > (in)Def PLURAL (APPROXIMATE)' where the definite plural usage of $liaŋ^{22}$ is directly developed from the numeral $liaŋ^{113}$, Shaoxing Wu '*liaŋ'* exhibits a distinctive grammaticalisation pathway: *liaŋ* 'two' > *liaŋ* 'two' + classifier > DUAL > (in)Def PLURAL (APPROXIMATIVE), which has not previously been mentioned in the literature.

6. Liaŋ^{22}-CL-NP: Evidence for the two-head hypothesis in Chinese DP

The question of the exact structure of the determiner phrase (DP), especially the sequence with the classifier and numeral in classifier languages, is debatable. In this section, we will discuss multiple possibilities to derive the syntactic structure of liaŋ^{22}-CL-NP under different assumptions about the DP sequences in Chinese. We suggest that the liaŋ^{22}-CL-NP construction can be perfectly derived if we follow the two-head approach to DPs in which the numeral and the classifier are assumed to instantiate two distinct heads, namely, Num and CL, respectively. In turn, the result provides fresh evidence for the two-head hypothesis.

6.1 Classifiers and numerals: Single-head or two-head?

Parallel to the verb phrase, the noun phrase is assumed to be mapped onto different functional projections, ultimately dominated by the determiner phrase (DP). Regarding the syntactic structure of DP sequences, however, no consensus has yet been achieved partially due to the complexity of the classifiers and numbers, especially in classifier languages. Two distinct hypotheses have been proposed regarding the syntactic status of classifiers and numeral forms: the single-head analysis and the two-head approach. In the single-head analysis, the classifier and number elements are assumed to form a single constituent projecting a NumP projection in which either the classifier hosts the head in the extended nominal projection (xNP); meanwhile, the numeral is a specifier of CLP or vice versa

(Tang,1990; Watanabe,2006; Naoki and Sakai, 2000 and others). In the two-head approach, the classifier and number are projected into two distinct functional projections where they instantiate their own functional heads, Num and CL (Pan,1990; Cheng and Sybesma,1999; Simpson,2005) or the classifier heads an individuation projection while the numeral serves as a specifier of #P or Num (quantity) (A. Li,1999; Borer, 2005). Some authors have argued further that different classifier languages have different DP structures depending on whether the classifier or the numeral appears independently (Saito et al., 2008; Hall, 2019). Overall, there has been much debate about the syntactic positions and semantic functions of classifiers and numbers in the literature. However, scant explicit argumentation with convincing empirical evidence has been raised to justify one of the possible analyses over another. Whether numbers and classifiers form a complex single head or instantiate two distinct heads is a question with significant consequences for linguistics. Therefore, arguments in favour of either of the two possible analyses should be explored further, as noted in Simpson (2005). In the following section, we will show that the data of 'lian^{22}-CL-NP' in Shaoxing Wu support a two-head hypothesis, or more precisely a Num-CL-two-head analysis for the Chinese DP.

6.2 Num-CL-two-head hypothesis

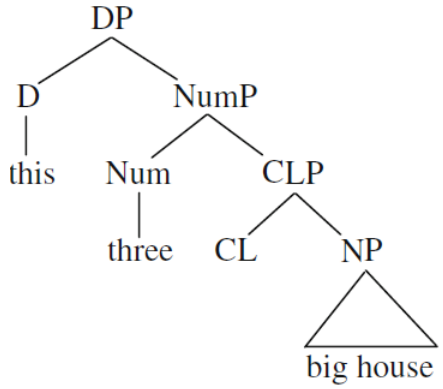
Similar to Pan's (1990) approach to Mandarin DP sequences, Simpson (2005) postulated a Num-CL-two-head DP structure and proposed a CL-to-D analysis to capture the cross-linguistic word order variations in DPs among Southeast Asian languages, such as Mandarin, Cantonese, Thai, Vietnamese, Hmong and Malay. For instance, in Mandarin and Thai, the DP sequences in these two classifier languages have distinct word orders on the surface, as abstracted in (55) below:

(55) **Mandarin DP:** [$_{DP}$ Dem Num CL Adj N]
Thai DP: [$_{DP}$ N Adj Num CL Dem]

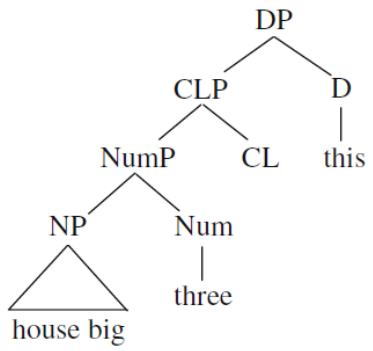
Considering that Thai is typologically a head-initial language while Mandarin behaves more like a head-final language, Simpson (2005) posited that the structures of Mandarin and Thai DPs should be mirror images of each other if the Num and CL form a single head, as claimed by the single-head hypothesis. Against the so-called single-head analysis, Simpson (2005) proposed a two-head structure for DPs, which is assumed to be a base-generated structure that is universally applicable to all classifier languages. One example is the Mandarin and Thai data from Saito et al. (2008).

(56) a. zhe san zuo da fangzi **(Mandarin)** b. baan yai saam lang nii **(Thai)**
this three CL big house house big three CL this
'these three big houses'

A.



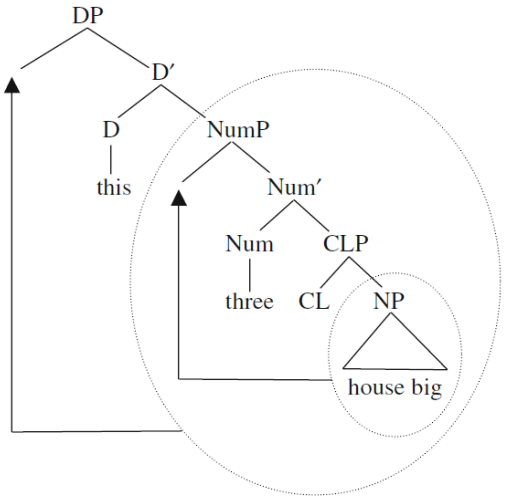
B.



.

According to Simpson, the Thai DP shares the base-generated structure with Mandarin, as in (56) A, and its distinct surface order, as in (56) B, can be derived by movement. Starting from (56) A, the NP moves to Spec of the NumP. Then, the NumP as a whole moves into DP Spec, which finally yields the surface word order in (56) B. The syntactic derivation of the Thai DP can be illustrated clearly in (57).

(57)



Simpson (2005) successfully accounts for the variation in DP sequences among East Asian languages by resorting to movement, while saying less on the motivation for such movement. In the current study, we focus on the syntactic position of numerals and classifiers and adopt the DP structure shown in (56) A to derive the (in)definite liaŋ^{22}-CL-NP sequence in Shaoxing Wu. This sequence suggests that the two-head hypothesis with the Num and CL instantiating distinct heads in their own projections is on the right track, at least in Shaoxing Wu.

6.3 'Liaŋ^{22}-CL-NP': Num is a head

As discussed in sections 3 and 4, the numeral 'liaŋ' in Shaoxing Wu is polysemic and can be interpreted as the number 'two' with a rising tone 113 and as an approximate with a level tone 22, denoting a nonspecific quantity. Similar to Mandarin, numerals in Shaoxing Wu co-occur with classifiers to denote quantity in most cases[6] in which the number and classifier have their own tone contours. In contrast, when used for approximation, the numeral 'liaŋ^{113}' changes its pitch into a level tone 22 and is semantically bleached. Meanwhile, the classifier loses its independent tone and

[6] According to Nomoto (2013), Japanese is a typical obligatory classifier language in which numeral modification of nouns without the assistance of a classifier normally is unacceptable. However, the classifier can be omitted in certain contexts where the number slot is filled by a nonspecific or large number denoting a vague quantity. For example:

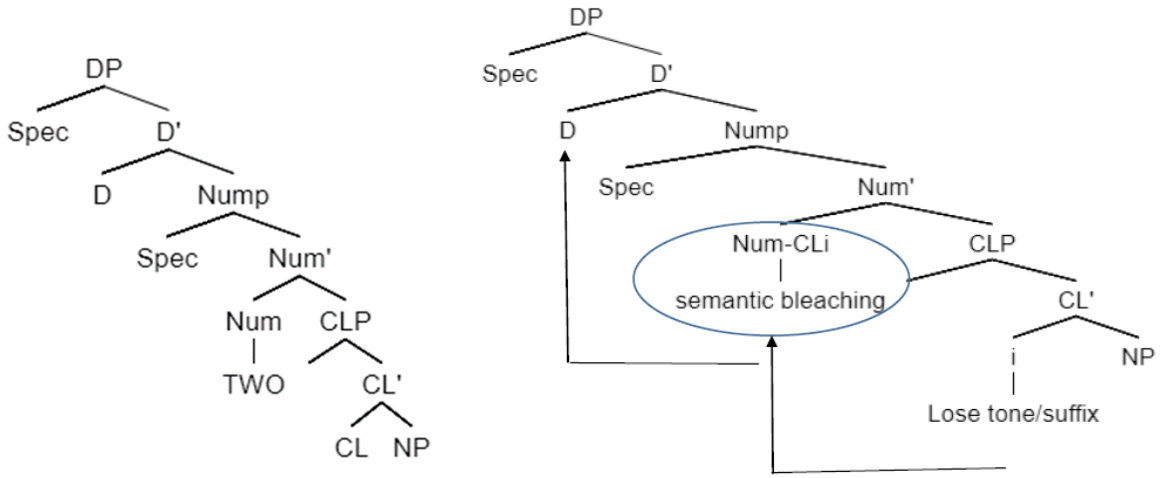
4. a. san *(ko)/kyuu? (ko)/juu-go (ko)-no gengo
three CLF nie CLF/fifteen CLF-LINK language
'three/nine/fifteen languages'
b. ni san (nin)-no gakusei
two three CLF-LINK student
'two or three/a few students'
c. juu-suu (ko)-no sima
ten-some CLF-LINK island
'more than a dozen islands'

Similarly, classifiers in Shaoxing Wu and Mandarin are optional when co-occurring with nonspecific or large numbers.

becomes like a suffix attaching to the preceding element, thus forming a phonological unit, or according to Feng (2005), a 'tonal unit' with 'lianŋ^{22}'. Moreover, the construction of 'lianŋ^{22}-CL-NP' can freely appear in the different definite contexts to represent plural definite reference. What is the syntactic structure of this new usage of 'lianŋ-CL-NP'?

Following Simpson's (2005) assumption that the number and classifier head their own projections, Num and CL, respectively, the structures of 'lianŋ^{113}(two)-CL-NP' can be assumed, as shown in (58)a. Then, the structure undergoes a grammaticalisation where the classifier loses its independent tone and becomes like a suffix attaching to the numeral. The numeral lianŋ with semantic bleaching constitutes a phonological unit with the classifier and moves further to the D position, where the classifier becomes phonologically null, as illustrated in (58)b.

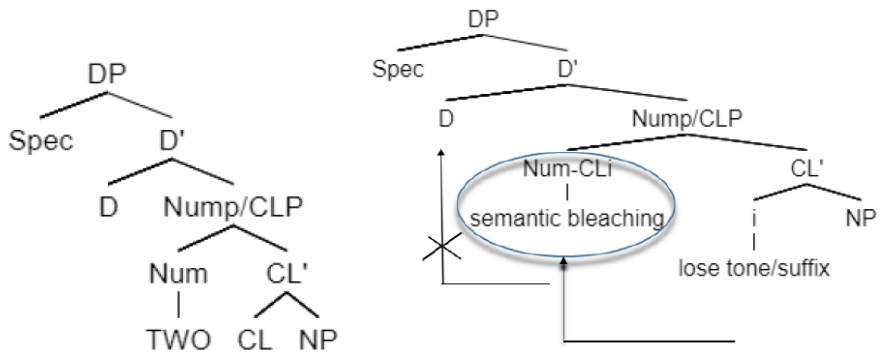
(58) a. quantity 'two' 'lianŋ^{113}-CL-NP' b. approximate 'lianŋ^{22}-CL-NP'



The DP sequence with the phonological realisation of D can be interpreted as definite or indefinite depending on the syntactic position where it finally appears. It is definite if it appears in the preverbal position in the sense of the VP external domain, as in Diesing (1992). Meanwhile, it is indefinite if it is within the VP domain and bound by existential closure. Such an assumption provides an account for the observation on NP interpretation in Shaoxing Wu in the literature, that is, preverbal NPs are definite, while postverbal ones are indefinite (Xu 2004; Yao 2018).

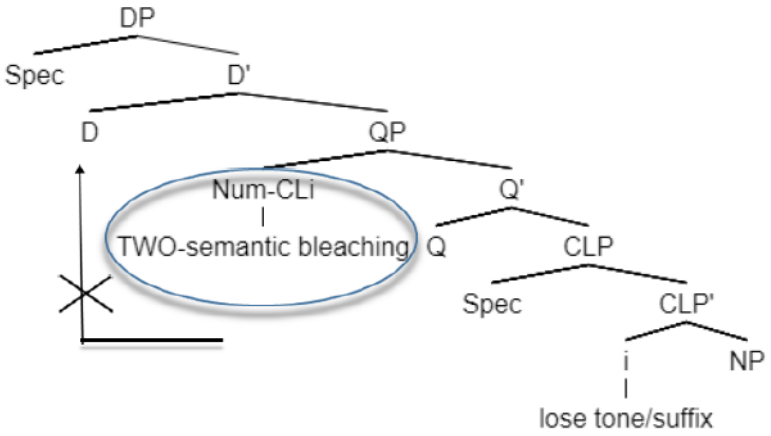
The classifier is normally assumed to be the head hosting a projection, while the numeral occupies the Spec (specifier) position wherever the projection is labelled CLP in the single-head analysis (Tang, 1990; Watanabe, 2006; Naoki and Sakai, 2000) or in some studies that support a two-head approach but still treat the numeral as a specifier (A. Li, 1999; Borer, 2005). We may attempt to derive the syntactic structure of the approximate (in)definite 'lianŋ^{22}-CL-NP' by assuming the numeral to be a specifier. The single-head hypothesis is considered first. A DP structure with one less projection is illustrated in (59) a.

(59) a. quantity 'two' 'lianŋ^{113}-CL-NP' b. approximate 'lianŋ^{22}-CL-NP'



Universally, it is assumed that D^0 must be overtly instantiated by some lexical element to trigger/signal a definite interpretation. Otherwise, the DP will be understood as having a default indefinite value (Simpson, 2005). Thus, the approximate Num-CL_i is motivated to raise to the D^0 position to obtain a definite reading if no other overt element appears in that position. In this configuration, however, the movement of the phonological unit of Num-CL_i is impossible because of an independently motivated ban on Head Movement out of a specifier (Roberts, 2010), as shown in (59)b. Moreover, such a movement will always be blocked as long as the numeral is analysed as a specifier in the DP, even if the numeral is located at another projection which is higher or lower than the one containing the classifier. For example, the numeral and classifier are projected into two distinct functional projections, as assumed in the two-head approach to the DP. The configuration in (60) below demonstrates the impossibility of the movement of 'Num-CL' to the D position.

(60) two-head hypothesis with the numeral as a specifier



Therefore, the impossibility of movement in configurations (59) and (60) suggests that the numeral and the classifier in Shaoxing Wu must host two different functional projections in order to guarantee that the approximate (in)definite 'lian^{22}-CL-NP' is derived correctly, as shown in (59)b above.

7. Conclusions

This study has examined in detail the properties of the approximative phrase 'lian^{22}-CL-NP' in Shaoxing Wu and its use as a definiteness device to represent plural reference. Partially following Simpson (2017), four different contexts are considered, a phenomenon not previously reported in the literature. Moreover, we investigated the alternations between 'lian^{22}-CL-NP' and other options to represent definite reference, such as bare CL-NPs, bare NPs and Dem-CL-NPs. Differing from

Mandarin and Jinyun Wu, Shaoxing Wu predominantly employs the 'liaŋ^{22}-CL-NP' construction to represent plural definite reference in cases of anaphoric definites, producer-product bridging reference and salient reference with or without visibility, and animate reference involving the speaker's personal relations. It alternates with the bare NP strategy in situations of part-whole bridging reference and inanimate reference, subject to the constraint of specificity. The more specific the definite reference is, the more the 'liaŋ^{22}-CL-NP' construction will be preferred.

Typologically, the use of 'liaŋ^{22}-CL-NP' in Shaoxing Wu as an approximate meaning 'a few-NP' to represent (in)definite plural reference suggests that it is undergoing a previously undescribed grammaticalisation process with the pathway of liaŋ 'TWO' + CLASSIFIER > DUAL > (in)Def PLURAL (APPROXIMATIVE). This grammaticalisation process, accompanied by the semantic bleaching of 'liaŋ̍ and loss of tone in the classifier, leads to a new usage of the 'liaŋ-CL-NP' construction, that is, from the typical function of denoting quantity to a definite device to represent plural reference. This usage also coincides with the syntactic structural changes of the DPs in Shaoxing Wu. These properties are captured by the two-head analysis, as proposed in (59) b.

Abbreviations

CL classifier
ASP aspect marker
PFV perfect aspect of verb
Prt particle
Q question marker
FSP sentential final particle
EXP experiential marker
Dem demonstrative
PL plural marker
DE relativizer
DU dual

ACKNOWLEDGEMENT

The work described was fully supported by a grant from the Research Grants Council of the Hong Kong Spec
Administrative Region, China (Project No. UGC/FDS24/H12/20).

Reference

Abbott, Barbara. 2004. Definiteness and indefiniteness. In Laurence R. Horn & Gregory Ward (eds.), Handbook of Pragmatics. Blackwell.

Aguilar-Guevara, Ana, Julia Pozas Loyo, and Violeta Vásquez-Rojas Maldonado. 2019. Definiteness across languages. Berlin: Language Science Press.

Barker, Chris. 2004. Possessive weak definites. In Ji-yung Kim, Yury A. Lander & Barbara H. Partee (eds.), Possessives and beyond: Semantics and syntax (University of Massachusetts Occasional Papers in Linguistics 29), 89–113. Amherst: GLSA.

Barwise, Jon and Robin Cooper. 1981. Generalized quantifiers and natural language. Linguistics and Philosophy 4, 159-220.
Borer, Hagit. 2005. Structuring sense: Vol. I: In name only. Oxford: Oxford University Press.
Bremmers David, Jianan Liu, Martijn van der Klis and Bert Le Bruyn. 2021. Translation mining: definiteness across languages (a reply to Jenks 2018). Linguistic Inquiry. Online early.
Chen, Ping. 2009. Identifiability and definiteness in Chinese. Linguistics 42(6), 1129-1184.
Cheng, Lisa Lai-Shen, and Rint Sybesma. 1999. Bare and not-so-bare nouns and the structure of NP. Linguistic Inquiry 30(4), 509-542.
Christophersen, Paul. 1939. The Articles: A Study of Their Theory and Use in English. Copenhagen: Munksgaard.
Dayal, Veneeta, and Julie Jiang. 2021. The puzzle of anaphoric bare nouns in Mandarin: a counterpoint to Index. Linguistic Inquiry. Online early.
Feng, Shengli. 2005. Hanyu Yunlu Yufa Yanjiu [Studies on Chinese phonological grammar]. Bejing: Peking University Press.
Gundel, Jeanette, Nancy Hedberg and Ron Zacharski. 1993. Cognitive status and the form of referring expressions in discourse. Language 69, 274–307.
Hawkins, John A. 1978. Definiteness and indefiniteness: A study in reference and grammaticality prediction. London: Croom Helm.
Hawkins, John A. 1991. On (in)definite articles: Implicatures and (un)grammaticality prediction. Journal of Linguistics 27, 405–442.
Heine, Bernd, and Tania Kuteva. 2002. World Lexicon of Grammaticalization. Cambridge University Press.
Heim, Irene. 1982. The semantics of definite and indefinite noun phrases. Amherst, MA: University of Massachusetts dissertation.
Himmelmann, Nikolaus P. 1996. Demonstratives in Narrative Discourse A Taxonomy of Universal Uses. In Studies in Anaphora, Edited by Barbara A. Fox, 205-54. Amsterdam: Benjamins.
Kuteva, Tania, Bernd Heine, Bo Hong, Haiping Long, Heiko Narrog and Seongha Rhee. 2019. World lexicon of grammaticalization (2nd edition). Cambridge University Press.
Lewis, David. 1979. Scorekeeping in a language game. Journal of Philosophical Logic 8, 339-359.
Li, Yen-Hui Audrey. 1999. Plurality in a classifier language. Journal of East Asian Linguistics 8, 75–99.
Li, Xuping, and Walter Bisang. 2012. Classifiers in Sinitic languages: from individuation to definiteness-marking. Lingua 122, 335-355.
Löbner, Sebastian. 2011. Concept types and determination. Journal of Semantics 28, 279–333.
Long, Haiping, Feng Gu, and Xiaoping Xiao (trans.), Bo Hong and Feng Gu (annotating). 2012. Translation of World Lexicon of Grammaticalization. Beijing: Beijing World Publishing Corporation.
Lü, Shuxiang. 1982. Aspects of Chinese grammar[Zhongguo Wenfa Yaolüe]. Beijing: commercial publisher. P: 365-366.
Lyons, Christopher. 1999. Definiteness. Cambridge: Cambridge University of Press.
Matthews, Stephen and Virginia Yip. 1994. Cantonese: a comprehensive grammar. London: Routledge.
Nomoto, Hiroki. 2013. Number in classifier languages. University of Minnesota dissertation.
Pan, Hai-Hua. 1990. Head Noun Movements in Chinese, the ECP and Minimality. Unpublished ms., University of Texas at Austin.
Peter Jenks. 2018. Articulated definiteness without articles. Linguistic Inquiry. Vol 49(3), 501-

536.
Roberts, Craige. 2003. Uniqueness in definite noun phrases. Linguistics and Philosophy 26. 287–350.
Roberts, Ian. 2010. Agreement and head movement: clitics, incorporation, and defective goals. (Linguistic Inquiry Monographs 59). Cambridge: MIT Press.
Russell, Bertrand. 1905. On denoting. Mind 14. 479–493.
Saito, Mamoru, T.-H. Jonah Lin and Keiko Murasugi. 2008. N'-ellipsis and the structure of noun phrases in Chinese and Japanese. Journal of East Asian Linguistics 17(3), 247–271.
Schwarz, Florian. 2009. Two types of definites in natural language. Amherst: University of Massachusetts. (Doctoral dissertation).
Simpson, Andrew. 2005. Classifiers and DP structure in Southeast Asia. In Guglielmo Cinque & Richard S. Kayne (eds.), Handbook of comparative syntax, 806–838. Oxford: Oxford University Press.
Simpson, Andrew. 2017. Bare classifier/noun alternations in the Jinyun (Wu) variety of Chinese and the encoding of definiteness. Linguistics. Vol 55(2), 305-331.
Sheng, Yimin. 2019. Quantitative interpretation and referential interpretation of Liang 'two' in the Shaoxing Wu dialect. Current research in Chinese linguistics. Vol 98 (2), 97-116.
Tang, Chih-Chen Jane. 1990. A note on the DP analysis of Chinese noun phrase. Linguistics 28, 337–354.
Wang, Li. 1944. Theory of Chinese Grammar. In Wang collection, Vol 1. Jinan: Shandong Education Publisher.
Wang, Xiaohui and Wang Xiaodong. 2012. Analysis of approximate usage of 'Liang X' in Mandarin. Language teaching and linguistic studies. Vol 6:81-88.
Watanabe, Akira. 2006. Functional projections of nominals in Japanese: Syntax of classifiers. Natural Language and Linguistic Theory 24(1), 241–306.
Xu, Liejiong. 2004. Manifestation of informational focus. Lingua. Vol 114 (3), 277-299.
Yao, Jennifer S.Y (2018). NP Interpretation and Disposal Variations among Mandarin, Cantonese and Shaoxing Dialect. Language and Linguistics. Vol 19(2), 306-332.