BARE-CLASSIFIER CONSTRUCTIONS IN XIANG (SINITIC): FROM TYPOLOGY TO FORMALIZATION*

Robert Marcelo Sevilla
The University of Hong Kong

ABSTRACT

The Sinitic 'bare-classifier' construction, where a noun follows a classifier without an accompanying numeral or demonstrative, demonstrates a complex interplay of syntax and semantics; associated with reference marking, the construction displays seemingly arbitrary restrictions on where it can occur and with what interpretation. This study introduces the variation of bare-classifier constructions among Xiang dialects, developing a typology of Xiang classifier constructions (drawing from typological work in Wang 2015) and formalizing this in a constraint-based OT framework. Syntactic and semantic constraints lead to the conclusion that bare-classifier constructions do not 'mark' reference, but rather restrict the possible referential interpretations available to the noun, while only directly marking individuation. The present paper explores bare-classifier constructions across four representative Xiang dialects: Changsha, Xiangxiang, Loudi, and Lianyuan, chosen as they represent a spectrum moving from full licensing of the bare-classifier construction in all contexts, to sole licensing of postverbal indefinites. Formalization of the observed variation following bidirectional OT syntax-semantics is shown to efficiently capture the typological trends observed.

Keywords: bare-classifiers, Xiang, typology, syntax-semantics, Optimality Theory

*I would like to acknowledge the help and support of my supervisor Prof. Stephen Matthews and my secondary supervisor Dr. Cathryn Donohue, both of whom provided indispensable comments and feedback on the piece. I would also like to thank two anonymous reviewers for their comments on the first iteration of this piece. Thanks should also go to Dr. Arthur Thompson for his helpful advice and comments on the draft.

1. INTRODUCTION

Classifiers (CLs) are a common feature of Sinitic languages (Aikhenvald 2000; Goddard 2005; Del Gobbo 2014), particularly known for noun classification and individuation, and generally occurring accompanying numerals, demonstratives, or quantifiers. Somewhat less well-known but no less studied, are those functions related to reference marking with the 'bare-classifier' construction, so-called because the classifier occurs with the noun unaccompanied by a numeral, demonstrative, or quantifier (Bisang 1999; Cheng & Sybesma 1999; 2003; 2014; Li 2013; Bisang & Wu 2017). For example, in Standard Mandarin (MD) one might encounter a string such as the following: *háizi kànle bù diànyǐng* child watch CL movie 'The child watched a movie', with the object showing a [CL-N] construction rather than the canonical [Num-CL-N], as in *yi bù diànyǐng* one CL movie 'a movie'. Most Sinitic varieties use this construction in contexts where the object NP has indefinite reference, and in others such as Cantonese (Matthews & Yip 2013), various Xiang varieties (Chen 1999; Wu 2005; He 2006), and many others (Wang 2015), [CL-N] constructions appear to serve as markers of definite and indefinite reference, or to at least be associated with changes in referential status of the modified NP.

In geographical terms, Southern Sinitic varieties tend to have more complex usage and functions of bare-classifier constructions in comparison to Northern varieties such as Mandarin, referred to by Chappell as an "extreme polyfunctionality of classifiers" (Chappell 2015: 48). For instance, Cantonese allows [CL-N] in both subject and object position and may mark definite and indefinite references (Cheng & Sybesma 2003; Matthews & Yip 2013). In comparison, Standard Mandarin does not license subject position [CL-N] and marks only indefinite reference in object position. These two varieties vary considerably from each other in the licensing of [CL-N], which is understandable given their wide divergences in other aspects of grammar (see Matthews & Yip 2013) and their physical distance. However, in varieties of the Xiang subgroup native to Hunan, which lie between these two large typological regions, the treatment of [CL-N] is comparatively

complex and relatively less studied, moderated as it is by syntactic, semantic, and discourse constraints.

Central Sinitic varieties such as Xiang, owing to influence from both Northern and Southern areas (Norman 1988; Wu 2005; Chappell 2015), demonstrate a combination of both systems, with some of the extensions found in the South in addition to many of the Northern limitations. In Xiang, this is evident not only when compared with other Sinitic languages, but even when one compares within Xiang, where one finds a high degree of variation in terms of permissible structures (Wu 2005; Wang 2015; Chappell 2015). Xiang dialects display referential usage of classifiers in various contexts; however, these are not found uniformly from dialect to dialect, presumably owing to different levels of contact with northern or southern Sinitic, which may have bleached native features (Norman 1988; Wu 2005). The variation and typology of classifier usage across the Xiang family is therefore poorly understood, and the realities of the distribution of reference-marking classifier constructions are understudied.

This state of affairs, in which classifiers seem to mark a diffuse referential category which is limited by positional licensing, necessitates further study. As an integral functional category of many numeral-classifier languages in East Asia and beyond (Aikhenvald 2000; Goddard 2005), as well as a compelling example of grammaticalization trends (Bisang 1999), research on these items is important for our understanding of how grammars work both through time and at present. The goal of the current paper is to explore the syntactic-semantic distribution of classifiers across four key Xiang dialects (based on extensive work in Wang 2015) and to formalize these in a constraint-based account, following an Optimality Theoretic approach (Prince & Smolensky 1993; 2004; Kager 1999; Kuhn 2003; Legendre et al. 2016). OT is particularly efficient at capturing typological generalizations with its theoretical machinery, as well as making predictions about the limits of said typology. Therefore, the present approach seeks not only to formalize the typological assumptions in Wang (2015) but also to theorize about the possible iterations of definiteness and syntactic positioning which may give rise to the different bare-classifier settings in Sinitic. It will be observed in the present study that the distribution and availability of Xiang bare-classifier

constructions are moderated by position relative to the verb and reference value, evidencing their quality of straddling the interface between syntax and semantics (Li 2013).

The present study adopts three key assumptions about the properties of referential use of bare-classifiers in Sinitic, further illustrated in §2. Firstly, classifiers are not fully grammaticalized as markers of definiteness (Li & Bisang 2012), and therefore do not *mark* any particular referential value (i.e. definite, indefinite) on the NP (at least not directly). The claim that bare-classifiers are markers of indefiniteness is based on their usage in Mandarin-type Sinitic languages (where bare-classifiers are often treated as null-numeral phrases; Cheng & Sybesma 1999), but in others it is rarely the case that classifiers unambiguously mark this semantic feature. Secondly, classifiers only unambiguously mark the semantic feature of individuation (atomization, countability), whereby an NP is viewed as composed of discrete entities. This is directly related to the more basic use of classifiers accompanying demonstratives, numerals, and quantifiers. Thirdly and finally, when it comes to definiteness, the function of bare-classifiers is to impose a restriction on the possible referential features of the NP (echoing arguments in LaPolla 2003), with a strong preference for the unmarked value in a certain position, rather than mark any particular value. There is evidence for these three claims from across Sinitic, found for the present study primarily in Wang (2015), which provides a typology of cross-Sinitic bare-classifier syntax-semantics.

Bare-classifier constructions vary in their appearance and interpretation across Xiang dialects (see Chen 1999; Liu 2001; Wu 2005; He 2007; Wang 2015), which is here believed to be determined by dialect-specific hierarchies of grammatical constraints. The current study will analyze a sample of four illustrative dialects, which represent the general typology of permissible Xiang classifier constructions, including the Changsha, Xiangxiang, Loudi, and Lianyuan dialects. According to Wang (2015), each of these varieties represents a different node on the spectrum of available bare-classifier constructions in Xiang, running from most to least restrictive in their grammatical and interpretational availability.

The remainder of the paper is organized as follows: §2 discusses the background of classifier usage and syntax in Sinitic, beginning with a descriptive account and an introduction to the typological framework of

Wang (2015) in §2.2, followed by a discussion of formal approaches to classifier syntax and referential features in §2.3 and §2.4. §3 introduces the Sinitic subgroup currently under analysis, Xiang, and takes a descriptive approach to classifier syntax-semantics in four Xiang dialects: Changsha, Xiangxiang, Loudi, and Lianyuan based on work in Wang (2015), He (2007), Liu (2001), and Chen (1999) (§3.1-6). This section outlines the data on which the theoretical analysis rests. §4 develops the constraint-based framework employed to analyze the variation in classifier constructions across these four Xiang dialects. Finally, §5 consists of a general discussion of the analysis and concludes with a discussion of limitations and possible future directions.

2. BACKGROUND: CLASSIFIERS AND REFERENCE IN SINITIC

The present section will outline some of the more basic facts about classifiers in Sinitic, starting with their defining features in §2.1, moving on to their use in referential constructions across Sinitic in §2.2. The section then addresses two previous approaches to classifiers and reference marking (§2.3), before concluding with a discussion of the extra-syntactic features explored in the analysis (§2.4).

2.1 Classifiers in Sinitic

Two primary features of classifiers include obligatory presence between numeral and noun phrases, and taxonomic division of nouns into semantic categories (Aikhenvald 2000; Goddard 2005; Del Gobbo 2014; Bisang & Wu 2017). The use of classifiers following numeral phrases is related to their function in marking 'individuation' (Bisang 1999; Cheng & Sybesma 2003; Li & Bisang 2012; Li 2013), whereby nouns are atomized into discrete entities, and can therefore be counted. Consider the following examples from Mandarin (provided in Pinyin):

(1) a. **yī* *táozi*
one peach

b. *yī* *ge* *táozi*
one CL peach
'One peach'
c. **sān* *zhū*
three pig
d. *sān* *tóu* *zhū*
three CL pig
'Three pigs'

Examples (1a) and (1c) demonstrate the ungrammaticality in Mandarin of a numeral modified NP lacking an accompanying classifier, while examples (1b) and (1d) demonstrate the use of specific classifiers based on the semantic class of the noun, in this case *ge* and *tóu*, corresponding to general objects and livestock, respectively. The numerals in the examples in (1) may all be replaced by either a demonstrative or quantifier, such as *zhè* 'this/these' or *měi* 'every', and would be equally ungrammatical without a classifier occurring between them and the noun.

In addition to these defining features, Sinitic classifiers are also related to NP referentiality (Bisang 1999; Cheng & Sybesma 2003; Li 2013; Wang 2015; Bisang & Wu 2017). The relationship of classifiers with definiteness arises from their more basic function of individuation, through a path of grammaticalization (Bisang 1999; Li & Bisang 2012). A definite or indefinite reading may arise in Sinitic languages with the 'bare-classifier construction' in the form [CL-N], such as in the following examples from Fuyang Wu (tones not provided in source) and Cantonese:

(2) a. *kɤ* *lɔpan* *ma-lə* *bu* *ts^{h}ots^{h}ɨ*
CL boss buy-PFV CL car
'The boss bought a car' (Bisang & Wu 2017)

b. *go3* *lou5baan2* *maai5-zo* *gaa3* *ce1*
CL boss buy-PFV CL car
'The boss bought a/the car' (Matthews & Yip 2013: 93)

Importantly, referential status of the classifier-modified NP varies based on its position relative to the verb; different Sinitic languages allow different interpretations depending on whether the [CL-N] sequence occurs pre- or post-verbally. This is hypothesized by Li & Bisang (2012) to be due to the incomplete grammaticalization of classifiers in Sinitic languages. Therefore, the classifier in [CL-N] is not a marker of definiteness per se (as, say, a determiner in languages like English), but rather functions to *restrict* the possible interpretations of the NP. In Fuyang Wu, seen above in (2a), preverbal bare-classifiers are interpreted as definite, while postverbal bare-classifiers are obligatorily indefinite (Li 2013: 245; Bisang & Wu 2017), while bare NPs would have a freer interpretation. Meanwhile, Cantonese in (2b) allows the same [CL-N] construction but with an ambiguous reference (either definite or indefinite based on discourse information) for the postverbal object. Compare this with the following examples from Standard Mandarin, which cannot have [CL-N] in preverbal position, and can only have indefinite readings of [CL-N] in object position:

(3) a. *(*gè)* *lǎobǎn* *mǎi-le* *(yī)* *(liàng)* *chē*
(*CL) boss buy-PFV (one) (CL) car
'The boss bought a car'

b. *lǎobǎn* *mǎi-le* *(zhè)* *liàng* *chē*
boss buy-PFV DEM CL car
'The boss bought (this)/the car'

The example in (3a) demonstrates the ungrammaticality in Mandarin of a preverbal bare-classifier, as well as optional use of the classifier *liàng* as a marker of indefiniteness (with or without the numeral *yī* 'one'). Meanwhile, the subject in both examples is interpreted as definite given its syntactic position, since that preverbal position in Mandarin tends to be occupied by topicalized NPs and that these are obligatorily definite (Li & Thompson 1981: 86; Li 2013: 119); an indefinite reading would require a ['one'-CL-N] construction (Xu 1997). However, *chē* is inherently ambiguous: while indefinites are preferred postverbally (i.e. tend to be objects or focused) in Mandarin (Chao 1968: 76; Xu 1997; Goddard 2005:

38; Li 2013), the object does not require overt marking by way of a classifier/numeral or demonstrative in order to be considered indefinite or definite, respectively, as seen in example (3b).

Many studies have recognized the importance of syntactic position on determining reference value in Sinitic (Chao 1968: 76; Xu 1997; Goddard 2005: 38; Li & Bisang 2012; Wang 2015). The subject/topic position is strongly correlated with definite reference, while the object/focus position is strongly correlated with indefinite reference. Information structure, as well as syntactic positioning, seems to also play a considerable role in determining the reference of [CL-N]; as others observe (Chafe 1987; Lambrecht 1994; Li & Bisang 2012), topic and focus are related to definite and indefinite reference, respectively. This helps explain why positioning, rather than marking of an NP with a bare-classifier, is a more causal factor in determining reference.

Although the vast majority of Sinitic languages allow [CL-N] in at least one context (see Wang 2015), they differ as to where it may occur and what reference value is assigned to the NP (moderated by syntactic positioning and information structure). Some varieties of Sinitic, for instance, disallow the occurrence of bare-classifiers in preverbal position or restrict their interpretation to exclusively definite or indefinite in pre- or postverbal position, respectively (Cheng & Sybesma 2003; Li 2013; Wang 2015; Bisang & Wu 2017). This cross-Sinitic variation is likewise reflected within the Xiang dialects on a smaller scale, which is efficiently captured in Wang's (2015) typology of bare-classifier constructions, discussed in the following section.

2.2 Typology of Referential Classifier Constructions

The syntactic distribution, availability, and semantic interpretation of bare classifier constructions is constrained by their position relative to the verb; that is, whether the [CL-N] phrase occurs pre- or post-verbally. This may surface as either a syntactic restriction on [CL-N] occurring pre- or post-verbally or as a semantic restriction on the type of reference which may be marked in a certain position. Wang (2015) has developed a series of implicational universals and seven types of permissible [CL-N] structures and interpretations found in Sinitic, through collected data from

120 Sinitic languages (11 of which are Xiang, the vast majority being Northern varieties). Wang proposes three implicational universals to account for the typological occurrence of the bare-classifier construction:

1. If a Sinitic language has preverbal [CL-N] it must also have postverbal [CL-N]
2. If a Sinitic language has preverbal indefinite [CL-N], it must also have definite [CL-N]
3. If a Sinitic language has postverbal definite [CL-N], it must also have indefinite [CL-N]

These implicational universals assume that postverbal [CL-N] will be more common than the preverbal equivalent, that preverbal [CL-N] will be definite more often where found, and that postverbal [CL-N] will be indefinite more often where found. Wang then uses these universal implications to develop a typology of bare classifier constructions and interpretations found across Sinitic; of interest to the current study are those types found in Xiang, of which there are four (from Wang 2015):

Type I: [CL-N] allowed with both Def/Indef readings pre- and post-verbally (Lianyuan, Xinhua, Xiangtan)
Type II: Definite [CL-N] excluded only post-verbally (Loudi, Ningxiang)
Type III: Indefinite [CL-N] excluded only pre-verbally (Xiangxiang)
Type VII: Only Indefinite [CL-N] in post-verbal position (Changsha, Shaoshan, Shaoyang, Changning, Hengdong, Qidong)

In addition to these four types, Wang lists three others which do not occur in Xiang: definite preverbal and indefinite postverbal [CL-N] only (Type IV), postverbal [CL-N] only with either interpretation (Type V), and no [CL-N] in either position (Type VI). It is interesting to note that even in the small sample of 11 Xiang dialects provided by Wang (2015), there is a wide degree of variation in terms of permissible [CL-N] occurrences, reflecting the high level of variation within this putative subgrouping. On the one hand, Type I dialects exhibit the full range of possible positions relative to the verb plus both possible reference types; on the other hand, we find Type VII languages (which, according to Wang, make up 90 out

of 120 surveyed dialects, with a strong Northern Sinitic bias) which allow only postverbal indefinite bare-classifiers (as in standard Mandarin).

Based on the above, three things are apparent: definiteness seems to be preferred pre-verbally, indefiniteness seems to be preferred post-verbally, and postverbal [CL-N] is preferred to preverbal [CL-N]. Wang explains this preference as due to two factors: increased phonological reduction in postverbal position, and the tendency for subjects to be definite and objects indefinite in Sinitic. Wang assumes that definite interpretations of [CL-N] originate from an underlying, phonologically reduced [DEM-CL-N] construction, while indefinite interpretations originate from an underlying ['one'-CL-N] construction. Following Bybee et al. (1990), Wang holds that phonological reduction tends to occur in postverbal position (nearer the end of an utterance, one would suppose), which would account for the overall preponderance of [CL-N] in that position, regardless of reference value, as each construction is reduced. Meanwhile, it is also observed that indefinite objects and definite subjects are treated as unmarked across Sinitic (Chao 1968: 76; Xu 1997; Goddard 2005: 38), which accounts for their preponderance in these positions.

Wang's analysis of DEM and 'one' reduction is suspect, since the distributional criteria *and* interpretation for [CL+N] is different than both [Dem-CL-N] and ['one'-CL-N] (Cheng & Sybesma 1999; Li & Bisang 2012; Li 2013: 248, 257; and §4). However, the typological tendencies in terms of syntactic positioning are well-founded, as mentioned in the preceding section. The latter issue is tied with the internal structure of the [CL-N] and how this relates to definiteness marking, covered in the following section.

2.3 Structure of the Bare-Classifier Phrase

This section presents two competing approaches to bare-classifier structure in Sinitic and concludes that the internal structure of the bare-classifier phrase should not be the main focus of its interpretational differences. Rather, semantic and pragmatic features such as individuation and familiarity, plus syntactic positioning, are believed to play a more fundamental role. Many Sinitic varieties share structural similarities;

however, as this and future sections will show, there are certain ways in which ClP syntax-semantics differs from language to language, particularly in ways relating to what semantic/referential features can be marked on the NP by means of a bare-classifier.

In accounts of classifier syntax in Sinitic, classifiers head their own phrases ClP, which are commonly thought to occur between the NP and NumP in a right branching structure (Cheng & Sybesma 1999; Hsieh 2008; Her & Hsieh 2010; Li 2013), all of which in turn occur within the DP. This structure is exemplified in the following, adopted from Her and Hsieh (2010) and Cheng and Sybesma (2014):

(4) [DP…[NumP…[ClP…[NP…]]]]

Assuming that this structure is equally applicable to Xiang varieties which display identical constructions, and using Changsha as an example, the following is adopted for DP phrases of the type *ko*^{24} *san*^{33} *tsa*^{24} *zən*^{13} 'these three people':

(5) [DP [D *ko*^{24}] [NumP [Num *san*^{33}] [ClP [Cl *tsa*^{24}] [NP [N *zən*^{13}]]]]]

Bare-classifier constructions were originally thought to be derived from a reduced numeral 'one', owing to referential similarities between ['one'-CL-N] and [CL-N] in that both are indefinite in Mandarin (Lü 1944, in Li & Bisang 2012). Cheng and Sybesma (1999) show, however, that there are distributional grounds for rejecting this hypothesis; in particular, they point out that [CL-N] constructions are limited to non-specific interpretations, while ['one'-CL-N] can be either specific or non-specific. In addition, they note that [CL-N] is specifically restricted from preverbal position in ways that ['one'-CL-N] is not. Comparison with other languages like Cantonese, which have definite [CL-N] (which is incompatible with a singular NumP), casts further doubt on this hypothesis.

They therefore posit that [CL-N] is semantically distinct and must have different structural representations corresponding to definite and indefinite reference: indefinite [CL-N] involving a numeral phrase as its maximal projection with an empty Num head, while definite [CL-N] has

a classifier phrase as its maximal projection. They propose the following two opposing structures:

(6) a. [NumP [Num ø] [ClP [Cl …] [NP [N …]]]] (Indefinite)

b. [ClP [Cl …] [NP [N …]]] (Definite)

The first structure in (6a) treats numeral phrases as inherently indefinite, and in turn holds that a null numeral is present in all indefinite readings of [CL-N]; the definite reading schematized in (6b) possesses a ClP as the maximal projection and is hypothesized to fulfill the role of the DP in other languages. In addition, in bare-nouns with definite reference, the CL head is held to be filled by the N. This effectively accounts for the interpretational differences through recourse to a null head.

This approach is unsatisfying for several reasons; primarily, there is the postulation of null heads lacking evidence that they are recoverable by native speakers (Li 2013). There is also the somewhat confusing rejection by Cheng and Sybesma of the reduced NumP hypothesis, which is then solved by positing a null NumP (Li & Bisang 2012). As an alternative, Li (2013: 259) proposes that the interpretational difference is reflected in a structural raising of Cl to D, proposing the structure in (6b) above as the *indefinite* construction, while movement of the classifier into the head of a DP projection accounts for the *definite* construction (with (6b) repeated below):

(7) a. [ClP [Cl …] [NP [N …]]] (Indefinite)

b. [DP [D Cl_i] [NumP [Num …] [ClP [Cl t_i] [NP [N …]]]]] (Definite)

As is observed in the above, step-wise head-to-head movement of Cl to D is held to account for the difference in referential status of the noun in (7b). This conforms with the approach to definiteness and the NP found in Lyons (1999) and Simpson (2005), where the DP is treated as a phrase encapsulating the marking of definiteness, rather than the word class of

determiners. Following this approach, any item expressing some value for definiteness is expected to fill the head position of the DP projection.

This approach also assumes similar movements into the DP in definite bare nouns, which would differentiate them from indefinites which have no projected DP. Therefore, the following would be postulated for definite (8a) and indefinite bare nouns (8b):

(8) a. [DP [D N_i] [NP [N t_i]]]

b. [NP [N …]]

Li's (2013) approach to Sinitic classifiers has several advantages over that proposed by Cheng and Sybesma (1999). Firstly, it correctly predicts that [Num-CL-N] constructions may be definite, given raising of the numeral into D. Secondly, given that personal names can be modified by classifiers in Wu dialects (Li 2013), it also accounts for the failure of personal names to raise from N to D when that position is already filled by a classifier or determiner, as in Italian or Spanish (Longobardi 1994). Finally, it allows us to do away with the problematic positing of a null numeral in order to account for indefinite NP reference, which has been the traditional approach to [CL-N] since Lü (1944). The major issue with this approach, however, is the postulation of covert movement, which is motivated by the checking of a functional feature. In addition, it unknowingly assumes that all definite NPs in Mandarin undergo covert movement, which, all else being equal, should be avoided on the grounds of parsimony.

Both of these approaches present problems, primarily due to the postulation of null elements and covert movements. Notably, the interaction between referential features and positional licensing could be relied on to explain the variation, rather than resorting to inaudible transformations. OT allows us to avoid positing movement operations to account for changes in bare-CL reference value; rather, we can assume that the position of a particular NP relative to the verb determines its referentiality, in addition to pragmatic considerations of familiarity-novelty. The features relevant to this account will be expanded upon in the following section.

2.4 Referential Features and Discourse Structure

The syntactic distribution of bare-classifier constructions is moderated by referential features determined by discourse structure; therefore, the grammaticality of [CL-N] constructions is constrained by their interpretation (Bisang 1999; Cheng & Sybesma 2003; Li 2013; Wang 2015; Bisang & Wu 2017). While this extends beyond previously discussed definiteness into related features such as specificity and genericity (Chierchia 1995; Li 2013; Cheng & Sybesma 2014), only definiteness will be discussed. Future research may show that additional semantic features are of paramount importance. However, for the time being, 'definiteness' as such will be considered as the primary factor determining the competition between bare NPs and bare ClPs in Xiang dialects.

The marking of definiteness has been explicitly linked with the availability and distribution of bare-classifiers in §2.2. 'Definiteness' will be treated as a feature of nouns indicating the identifiability or familiarity of entities in the discourse context (Lyons 1999: 5, 6; Li & Bisang 2012, Li 2013: 121). NPs which satisfy the condition of being identifiable or familiar in the discourse context will be treated as 'definite'. In turn, NPs which do not satisfy this condition, being novel to the discourse context, will be treated as 'indefinite'. The contrast is captured through binary values for discourse or pragmatic information (PI) relating to familiarity: [+Fam] and [-Fam]. The current definition of definiteness agrees with Li & Bisang (2012) in that in Sinitic, this category corresponds primarily to the pragmatic notion of familiarity, rather than simply uniqueness or existence (Sio 2007). At the level of interpretation, definiteness as a collection of features (not just familiarity) is marked through the binary feature [Def], which is believed to go beyond simple familiarity.

A core argument made here is that classifiers as such do not directly *mark* definiteness or NP reference, but rather that they only directly mark individuation of the NP (that is, whether the referent of the NP is composed of discrete entities, primarily for purposes of countability). In fact, there seems to be a historical relationship between the individuating function and referential functions of classifiers (Bisang 1999; Li & Bisang

2012). Synchronically, definiteness or reference value is here held to be picked up instead from syntactic position and only indirectly through the presence of a classifier; therefore, the presence of the classifier is assumed to mark an individuated NP, carrying a functional individuation feature: [+/-Ind].

While the classifier is assumed to only directly mark individuation, there is clearly a relationship with definiteness: the bare-classifier imposes a restriction on the possible referential values of the NP (with a preference for unmarked values in each position). Therefore, other intervening factors serve to directly provide a particular NP referential value, including position relative to the verb and pragmatic information such as familiarity-novelty. In MD, for example, the [CL-N] restricts the interpretation of the postverbal NP to non-specific singular, in contrast to both the ['one'-CL-N] or bare-noun, which may be specific or non-specific (for the NumP), or definite, indefinite (spec. or non-spec.), plural, singular, or generic (for the NP) (Cheng & Sybesma 1999; Li & Bisang 2012). One would not argue, however, that the classifier 'marks' non-specificity since this interpretation is clearly still available in the other two constructions. This argument reflects that found in LaPolla (2003), who sees increasing linguistic structure as functioning to constrain interpretational contexts for the listener, thereby facilitating discourse. Xiang dialects differ in terms of what referential values they license in a particular position; this is here considered evidence of this restrictive quality, as well as the hierarchical ranking of formal constraints which either prefer or discourage the interpretation of certain referential values by position and disallow the presence of the [CL-N] construction in preverbal positions. The status of bare-classifier constructions in four representative Xiang dialects is explored in the following section.

3. BARE-CLASSIFIERS IN XIANG

Xiang is a family of Sinitic languages spoken in Hunan province by around 30 million people (Ethnologue 2018: Chinese, Xiang). In typological terms, it occupies an interesting place within Sinitic, as its geographical position between the typologically Northern and Southern

Sinitic languages has given Xiang a 'transitional' grammatical profile (Norman 1988; Ramsey 1989: 97; Wu 1999, 2005; Chappell 2015). It is for this reason that many of the features of classifiers considered characteristic of Southern Sinitic (definite bare classifier constructions, preverbal bare classifier constructions, etc.) are found primarily in more conservative Xiang dialects (which tend to cluster in central and southern Hunan). However, these are not uniformly distributed even within these dialects, and there is therefore a high degree of variation as well.

The greater degree of variation and conservatism observed for central Xiang dialects means that they represent three key types in Wang's survey (§2.2). The vast majority of Xiang dialects represent Type VII (the most common type across Sinitic), and therefore the most extensively studied dialect, that of Changsha, was selected to represent this category. In order to represent the remaining possibilities in bare-classifier constructions, the dialects of Xiangxiang (III), Loudi (II), and Lianyuan (I), are selected owing to their relatively well-documented status.

The focus of this study is on four dialects representative of the different bare-classifier constructions found in Xiang: Changsha, Xiangxiang, Loudi, and Lianyuan. Each of these represents a step from more to least complex in terms of permissible syntactic structures and interpretations available for classifiers in the language, excepting Xiangxiang and Loudi, which simply represent opposite configurations without any increase in complexity. The following sections will introduce each of the dialects under consideration and will also provide illustrations of their usage of the bare-classifier construction in reference marking.

3.1 Changsha

The speech of the city of Changsha (CS), located in north-eastern Hunan, represents the dialect of the provincial capital and is thus often treated as representative of other Xiang dialects (Chappell 2001; c.f. Wu 2005). However, this is a rather misleading characterization; Changsha is a particularly innovative variety of Xiang which tends to pattern grammatically with Northern Sinitic languages such as Mandarin (Norman 1988; Chappell 2001). Due to considerable historical and ongoing contact with Northern Sinitic, it is no surprise that the dialect

patterns more closely with this group in terms of its classifier syntax than with other Xiang dialects.

As mentioned in §2.1, Mandarin and the majority of Sinitic varieties only license [CL-N] in postverbal position with an indefinite reading only. This is also the case in CS; consider the following examples (author's data):

(9) a. *(*tsa^{24})* *lau^{13}pan^{13}* *mai^{41}-ta^{21}* *pu^{41}* *ts^{h}ɤ^{33}*
(*CL) boss buy-PFV CL car
'The boss bought a car'

b. *(*tsa^{33})* *tɕy^{33}* *li^{41}* *wei^{45}-lɤ* *ma*
(*CL) pig 2.SG feed-PFV Q
'Have you fed the pig?'

c. *li^{41}* *wei^{45}-lɤ* *tsa^{33}* *tɕy^{33}* *ma*
2.SG feed-PFV CL pig Q
'Have you fed a pig?'

d. *li^{41}* *wei^{45}-lɤ* *la^{45}* *tsa^{33}* *tɕy^{33}* *ma*
2.SG feed-PFV DEM CL pig Q
'Have you fed that pig?'

As examples (9a) and (9b) illustrate, preverbal [CL-N] is ungrammatical in Changsha, even if the classifier is attached to a topicalized object (9b). Postverbal [CL-N] is grammatical with an indefinite reading, as in *pu^{41} ts^{h}ɤ^{33}* 'a car' (9a) and *tsa^{33} tɕy^{33}* 'pig' (9c). This configuration is identical to that found in Standard Mandarin (c.f. §2). A definite interpretation can be achieved for the postverbal object if a demonstrative is used (as in 9d), while for a bare classifier to occur in preverbal position, it must be accompanied by a numeral, demonstrative, or quantifier, as in Standard Mandarin. In terms of Wang's (2015) typology of classifier references (c.f. §2.2), this would place Changsha in Type VII, along with the vast majority of Sinitic languages in the survey.

3.2 Xiangxiang

The Xiangxiang dialect (XX) is spoken to the southwest of Changsha in central Hunan, nearer to Loudi and Lianyuan. Following data in He

(2006), which describes the Xiangxiang spoken in Hutian, it can be observed that the dialect allows definite interpretations of [CL-N] in both post- and preverbal position, but specifically disallows indefinite only in the preverbal position. However, there is no ambiguity in terms of reference in the postverbal position, where either an indefinite or definite interpretation is possible, because the difference is coded in terms of tone change (Wang 2015). Consider the following examples:

(10) a. *tɕjɔ$^{33/13}$ tɕjɤ^{42} tsaw$^{55/35}$ mɤ^{11}tsɿ kʰaŋ$^{33/13}$ xetsʰɿ^{11} kɤ ŋɤ^{42} le*
CL dog do Q bark always ?? NEG stop
'Why is the dog always barking non-stop?'

b. *ɳ$^{11/42}$ tɕʰi$^{55/35}$ kʰuɛ̀^{35} kʰɔ$^{11/42}$ tɕjɔ$^{13/33}$ njɔ^{13} ta*
2.SG go see down CL ox SFP
'Go check on the/an ox' (He 2006: 12, 13)

Example (10a) demonstrates the obligatorily definite interpretation of the noun in preverbal position; an indefinite interpretation cannot occur in this location. Example (10b) shows that the postverbal [CL-N] *tɕjɔ$^{13/33}$ njɔ^{13}*, where both interpretations are possible, may be disambiguated through tonal change; the general classifier *tɕjɔ^{13}* with a rising tone indicates indefiniteness, while a mid-level tone form *tɕjɔ^{33}* indicates definite reference. Regardless, the fact remains that either interpretation is available to bare-classifiers in postverbal position, while the indefinite is disallowed in preverbal position. This setting corresponds to Wang's Type III typological classification, which is shared with a total of eight other dialects in the survey, including many varieties of Yue (such as Cantonese).

3.3 Loudi

The Loudi dialect (LD) is spoken near the Xiangxiang and Lianyuan localities in Central Hunan. Liu (2001) shows that the dialect allows the opposite configuration as that found in Xiangxiang, with only postverbal definite interpretations disallowed, as opposed to preverbal indefinites in Xiangxiang. Consider the following:

(11) a. *ko*^{42} *to*^{33}*-pa* *to*^{33} *ko* *ɛy*^{44}, *tɕjɔ*^{13} *siɔŋ*$^{44/33}$*-tsɿ* *p*^{h}*ɔ*$^{35/55}$
this many-AFF many ASP book, CL box-NOM afraid
xɔŋ$^{35/55}$ *p*^{h}*u*$^{35/55}$*lo*^{35} *pa*
put NEG fall SFP
'(I'm) afraid that the/a box cannot contain so many books'
(Wang 2015)

b. *wẽ*^{42} *mi*$^{42/11}$ *jɔŋ*$^{42/11}$ *ko* *ŋẽ*^{44} *nin*$^{13/33}$ *tã*^{45} *mi*$^{42/11}$ *jɔŋ*$^{42/11}$ *ko*
bowl rice raise CL kind person *dan* rice raise CL
zjo^{13} *nin*$^{13/33}$
enemy person
'A bowl of rice produces a benefactor, one *dan* (50kg) of rice
produces an enemy.' (Liu 2001: 297)

In the first example, the preverbal ClP *tɕjɔ*$^{13/33}$ *siɔŋ*$^{44/33}$*-tsɿ* 'the/one box' is ambiguous in terms of its interpretation, which may be either definite or indefinite depending on discourse context; i.e. if for instance there is a specific box present which is being referred to at the time the utterance is made. In the second example the two postverbal objects *ko ŋẽ*^{44} *nin*$^{13/33}$ 'a benefactor' and *ko zjo*^{13} *nin*$^{13/33}$ 'an enemy' are inherently indefinite. This conforms to Wang's Type II profile, which in his sample is limited to two dialects within the Xiang subgroup.

3.4 Lianyuan

The Lianyuan dialect (LY) is of interest to the current study for its unique place within the typological framework devised by Wang (2015): it occupies the opposite extreme as Changsha in terms of permissible [CL-N] interpretations, allowing both possible interpretations in both positions (Chen 1999). This puts it at one end of the spectrum, which is occupied at the other end by Changsha, which has the most limitations in terms of permissible [CL-N] structures and interpretations (discounting certain Min dialects; see Cheng & Sybesma 2005). Consider the following examples:

(12) a. *ko*^{55} *ɳen*^{13} *iə*^{55} *to*^{44} *tsau*^{55} *ti*^{55} *xə*^{42} *sʅ*^{11} *tɕi*
CL person should more do CL good things SFP
'A/the person should do more good things'

b. *ɳ*^{42} *tɕ*^{h}*i*^{55} *k*^{h}*uɛ*^{55} *xɔ*^{11} *ʈɔ*^{33} *ɳau*^{13} *ʈa*
2.SG go see down CL ox SFP
'Go check on the ox.' (Chen 1999: 274-5)

In example (12a) the subject *ko*^{55} *ɳen*^{13} 'person' is ambiguous; it may either be definite or indefinite. The gloss may be misleading; the discourse context implies an indefinite reading, but ostensibly either interpretation is possible. Meanwhile, the same holds of the object *ti*^{55} *xə*^{42} *sʅ*^{11}'good things'. While the same should hold in theory in example (12b) below, *ʈɔ*^{33} *ɳau*^{13} 'ox' is not ambiguous thanks to discourse information such as familiarity; this fact will be of paramount importance in the formalization in §4. This corresponds to Type I in Wang's terminology, in which a dialect can have definite or indefinite [CL-N] in both pre- and postverbal position.

3.5 Conclusions

The above examples illustrate the variation in terms of permissible structures found across these four Xiang dialects, which can be conceptualized as a spectrum from most to least constrained use of bare-classifiers as markers of reference. On one end of the spectrum, Changsha does not allow preverbal [CL-N], allowing it only in postverbal position with a strictly indefinite interpretation. On the opposite end, Lianyuan allows [CL-N] both pre- and post-verbally with both definite and indefinite interpretations. Meanwhile, in between these two are Xiangxiang, which has no preverbal indefinites, and Loudi, which has no postverbal definites.

What is interesting about this variation is how the constraints on bare-classifier phrases interact with the more common Sinitic reference marking through syntactic positioning mentioned by Wang in his analysis. This is why we have a change of reference in Mandarin examples such as:

(13) a. *shù shàng yǒu xiǎo niǎo*
tree above have small bird
'There is *a* small bird in the tree' (Not: 'The small bird')
b. *xiǎo niǎo zài shù shàng*
small bird be.at tree above
'*The* small bird is in the tree' (Not: 'A small bird')

The referentiality of the NP [*xiao niao*] 'small bird' changes based on whether it is topicalized, therefore preceding the verb. Additionally, it is interesting to note that the grammaticality of these constructions is dependent both on syntax and on semantics, in that bare classifiers are barred entirely from occurring in certain positions and are constrained in their interpretation when appearing in others. The competition between bare NPs and bare ClPs is dependent on two linguistic levels. This lends itself to an OT approach which takes account of both production (the syntactic component) and interpretation (the semantic component), as outlined in the next section.

4. BARE-CLASSIFIERS AND (BIDIRECTIONAL) OPTIMALITY THEORY

As mentioned in the preceding section, it is here held that there are considerable theoretical gains to be made through reformulating traditional generative analyses of [CL-N] syntax-semantics in terms of OT constraints. For one, it allows us to do away with claims of covert movement and trace elements. For another, it allows us to avoid positing null numeral and demonstrative elements as is done in Cheng and Sybesma (1999) and Wang (2015). Meanwhile, in addition to benefits in terms of formal syntactic parsimony, one of the major advantages of Optimality Theory lies in its ability to effectively formalize and account for typological variation (Kager 1999; Kuhn 2003), variation which is clearly seen in Wang's (2015) survey of Sinitic classifier constructions.

In developing a constraint-based analysis of bare-classifier constructions, two approaches to OT syntax-semantics are necessary to account for the Xiang data. The first is essentially the standard

unidirectional approach to OT syntax, which draws its essential characteristics from foundational work in Legendre et al. (1995) and Grimshaw (1997), and further work in Kuhn (2003) and Legendre (2001). It also borrows partly from work in Choi (2001) in its use of binary input features and pragmatic information. The second approach is the unidirectional approach to semantics, borrowing from de Hoop and Lamers (2006) and Legendre et al. (2016). In particular, this second approach is necessary to account for facts about XX, LD, and LY, all of which demonstrate optimization applying at the level of meaning; it is also important in stressing the fundamental (yet somewhat arbitrary) differences between CS and the other dialects. These two approaches coalesce in a sequential bidirectional OT approach outlined in Kuhn (2003), originally formulated in Blutner (2000), but without the formal machinery. This final expansion is ultimately necessary in order to account for the CS data, which must optimize first at the syntactic level, and then at the semantic.

The remainder of this section is organized as follows: §4.1 discusses the unidirectional approach to syntactic form, with an introduction to the production input and postulated constraints. This section also demonstrates the asymmetry between Changsha and the other three dialects. §4.2 provides the unidirectional approach to meaning optimization, accounting for the Xiangxiang, Loudi, and Lianyuan data. This section concludes with a brief introduction of the sequential bidirectional optimization approach, which is ultimately necessary to account for the Changsha data.

4.1 Unidirectional Form Optimization

In the standard approach to OT syntax, optimization is unidirectional in that only one level, be it production or perception, is optimized at a time. Output candidates are generated in response to an input, which are then evaluated against each other in terms of hierarchically ranked constraints. Candidates which violate more highly ranked constraints are eliminated; the one which has the fewest violations, or violates only low-ranked constraints is selected as the output and is therefore ‘optimal’.

The production input in the standard account is generally composed of a predicate, its arguments, and any functional categories or features (Grimshaw 1997; Legendre 2001). Functional features relevant to the competition of bare-noun vs. bare-classifier constructions include [+/-Ind], determining whether the NP referent is a discrete individual entity or not. In addition, it is assumed that some specification for discourse or pragmatic information is present as well, namely [+/-Fam], although this only comes into effect in the later comprehension stage (see §4.2). Consider the following input, corresponding to the CS output for 'The boss bought a car':

mai^{41} 'buy' (x, y)
x = *lau*^{13}*pan*^{13} 'boss' [+Ind, +Fam]
y = *ts*^{h}*ɤ*^{33} 'car' [+Ind, -Fam]
ASP = PFV
NC = General (x), Vehicle (y)

The above demonstrates the assumed input for a simple proposition with a two-place predicate: the predicate 'buy', two lexical heads (x) and (y), and two functional heads specifying tense/aspect and noun class. 'Noun Class' in Sinitic is a freer category than the name might suggest; here it is included simply as a method to determine what classifier *may* occupy the ClP head. The key point here is that classifiers are functional heads (following discussion in Packard 2000) and are therefore not input specified. The object *ts*^{h}*ɤ*^{33} is [+Ind], licensing the CL-N construction in CS since it occurs postverbally *and* since classifiers are held to mark individuation. In addition, both arguments are specified for discourse information in terms of familiarity. For the current application, focus is placed solely on the two NPs and the verbal properties of the predicate are largely ignored as only two-place predicates will be analyzed.

As mentioned in preceding sections, the presence of reference marking bare-classifier constructions in Xiang is constrained by position relative to the verb (post- v. preverbal, subject v. object, topic v. focus, etc.) and by interpretation (definite v. indefinite). Two related markedness hierarchies can be established building on the universal implicatures found in Wang

(2015) (where '>' means 'less marked than' or 'is implied by the presence of'):

postverbal [CL-N] > preverbal [CL-N]
postverbal-INDEF > postverbal-DEF
preverbal-DEF > preverbal-INDEF

These three hierarchies capture three generalizations about Sinitic languages: [CL-N] is universally marked in preverbal as opposed to postverbal position, postverbal definite [CL-N] is comparatively more marked than postverbal indefinite [CL-N], and preverbal indefinite [CL-N] is comparatively more marked than preverbal definite [CL-N].

In pursuing an OT account of the data, a series of constraints must be posited mediating semantic interpretations and syntactic positioning of [CL-N] constructions, based on the markedness observations seen in Wang (2015). The constraints are divided into two groups, representing a conflict between Markedness, or a general ban on 'marked' structures, and Faithfulness, or the requirement that input features be overtly marked in the output. Alternatively, Faithfulness can be understood as the requirement that everything in the input be parsed.

Given the wide preponderance of dialects which limit the presence of preverbal [CL-N] independent of interpretation, it is sensible to assume that this is a marked configuration. We therefore might posit the following constraint which penalizes their presence in the output:

(C1) *PRE-CL: no preverbal [CL-N]

No variety could be identified in Wang's survey which unilaterally banned postverbal [CL-N] but allowed preverbal [CL-N]; that is, the presence of preverbal bare classifier constructions implies the presence of postverbal [CL-N] in every case. In addition, the ban on bare-classifiers in this position seems to be unmotivated by any feature checking or arbitrary restriction (Li & Bisang 2012). Therefore, it is posited that there is such a constraint operating on the appearance of preverbal bare-classifiers.

While the vast majority of Sinitic languages allow the presence of bare-classifiers, there is a significant minority which ban them outright, particularly certain Min varieties such as Hokkien (Chen 1958; Zhou 1991; Cheng & Sybesma 2005). This provides the motivation for the following outright ban on their appearance, which is the logical equivalent of (C1):

(C2) *POS-CL: no postverbal [CL-N]

This constraint does not have any overt effect on the presence of bare-classifiers in Xiang dialects; its function is subsumed under *PRE-CL. It is assumed to exist given facts about Sinitic typology, such as the ban on [CL-N] in certain Min varieties just mentioned above. These two markedness constraints are opposed by the following faithfulness constraint, motivated by the need to mark individuation of the NP referent through a classifier:

(C3) FAITHIND: a value for [Ind] should be reflected in the output

This constraint finds its motivation in that nouns in the bare-classifier construction are unable to have a generic or 'kind' interpretation (Cheng & Sybesma 1999), and that generic classifiers with kind interpretations cannot occur in the [CL-N] construction. For instance, generic classifiers in Cantonese (Matthews & Yip 2013), which do not individuate entities, cannot occur in the [CL-N] construction without a demonstrative (Stephen Matthews p.c.); e.g. **(li) zung2 jan4* DEM CL.type person 'This type of person'. As mentioned in §2, a fundamental function of classifiers is the atomization of nouns for purposes of countability (Bisang 1999; Li & Bisang 2012).

The candidate set for the input should include all possible iterations of bare-noun and bare-classifier constructions, all of which correspond to possible Xiang outputs. For instance, consider the following four logical candidates in (14), followed by their Changsha equivalents in (15) for illustrative purposes:

(14) Candidates:
a. [CL-N]_V_[CL-N]
b. [CL-N]_V_[N]
c. [N]_V_[CL-N]
d. [N]_V_[N]

(15) CS Equivalents:
a. *tsa^{24} $lau^{13}pan^{13}$ mai^{41}-ta^{21} pu^{41} $ts^{h}ɤ^{33}$
b. *tsa^{24} $lau^{13}pan^{13}$ mai^{41}-ta^{21} $ts^{h}ɤ^{33}$
c. $lau^{13}pan^{13}$ mai^{41}-ta^{21} pu^{41} $ts^{h}ɤ^{33}$
d. $lau^{13}pan^{13}$ mai^{41}-ta^{21} $ts^{h}ɤ^{33}$

The structures in (14) differ in terms of the presence of bare-noun and bare-classifier constructions; in CS, only (c) and (d) are grammatical. However, in cases where the object checks for a positive [Ind] feature, candidate (c) will be preferred. At this point, we have enough theoretical machinery to account for the CS data in §3 (at least at the syntactic level), with the following ranking:

CS: *PRE-CL >> FAITHIND >> *POS-CL

CS allows only postverbal bare-classifiers to surface, which demonstrates the crucial ranking *PRE-CL >> FAITHIND (as this unilaterally bans preverbal [CL-N]); at the same time, the presence of postverbal [CL-N] shows that FAITHIND >> *POS-CL, since if the opposite ranking held, no bare classifiers would surface. Consider the following:

Table 1. Unidirectional approach to CS productive output

x = *lau*^{13}*pan*^{13} [+Ind, ~Fam] y = *ts*^{h}ɤ^{33} [+Ind, ~Fam] CL = General(x), Vehicle(y)	*PRE-CL	FAITHIND	*POS-CL
a. [*tsa*^{24} [*lau*^{13}*pan*^{13}]] _ [*pu*^{41} [*ts*^{h}ɤ^{33}]]	*!		*
b. [*tsa*^{24} [*lau*^{13}*pan*^{13}]] _ [*ts*^{h}ɤ^{33}]	*!	*	
→ c. [*lau*^{13}*pan*^{13}] _ [*pu*^{41} [*ts*^{h}ɤ^{33}]]		*	*
d. [*lau*^{13}*pan*^{13}] _ [*ts*^{h}ɤ^{33}]		**!	

As can be observed from the table above, candidates (a-b) both fail to surface as outputs due to their violations of the highly ranked *PRE-CL constraint, which penalizes the appearance of bare-classifiers preverbally. Meanwhile, candidate (d) fails due to its double violation of the

FAITHIND constraint, which requires [+Ind] input NPs to be classifier-modified. This provides the correct output in candidate (c). Notice also that for the time being discourse information is unspecified in production, as this feature is not directly marked through the [CL-N] construction; it will have a role in determining optimal comprehension in the following section.

If the input preverbal NP has a negative value for [Ind], the same output will hold, since a candidate without a preverbal [CL-N] will always be optimal in CS (owing to the high ranking of *PRE-CL). However, if the postverbal NP has a negative value, candidate (d) will surface as the output; this is due to the dispreference for [CL-N] when a more economical alternative is available (owing to the effect of *POS-CL surfacing when there is no feature to check).

While the constructions in (1a-b) are ungrammatical in CS, in dialects like XX, LD, and LY all four constructions are available. Therefore, we would posit that the following ranking holds in all three, at least at the syntactic level:

XX, LD, LY: FAITHIND >> *POS-CL, *PRE-CL

It is assumed that marking of the [Ind] feature is of primary importance in these dialects, motivating the ranking of FAITHIND above both other constraints. This interaction is observed in the following table, using LY input for ‘A person should do more good things’:

Table 2. Unidirectional approach to LY productive output

x = *ŋen*^{13} [+Ind, ~Fam] y = *xə*^{42} *sɿ*^{12} [+Ind, ~Fam] CL = General(x), Abstract?(y)	FAITHIND	*PRE-CL	*POS-CL
→ a. [*ko*^{55} [*ŋen*^{13}]] _ [*ti*^{55} [[*xə*^{42}] *sɿ*^{12}]]		*	*
b. [*ko*^{55} [*ŋen*^{13}]] _ [[*xə*^{42}] *sɿ*^{11}]	*!	*	
c. [*ŋen*^{13}] _ [*ti*^{55} [[*xə*^{42}] *sɿ*^{11}]]	*!		*

d. [*ɳen*^{13}] _ [[*xə*^{42}] *sʅ*^{11}]	**!		

This table shows the proposed productive output for XX, LD, and LY; essentially, if both input NPs are specified as [+Ind], a bare-classifier phrase can occur in any position owing to the highly ranked FAITHIND constraint. If either of them has a negative value for this feature, one of the lower ranked constraints will favor a bare-NP as an output.

This basic approach can account for the syntactic facts in question, such as the presence or absence of preverbal [CL-N]. However, this approach fails to account for the limitations on the possible referential interpretations of [CL-N] in all four dialects; for instance, the postverbal [CL-N] in CS example (2c) is obligatorily indefinite, and XX and LD limit the appearance of [CL-N] when the reference is definite (object) or indefinite (subject), respectively. To derive these facts, an additional complication is required, acting on the perceptual-semantic level, rather than merely the syntactic.

4.2 Unidirectional Meaning Optimization

In the OT approach to meaning optimization, the input is equal to the output of form optimization; that is, the perception input is a linguistic string corresponding to the winning production candidate. However, assuming that no string occurs in isolation, there must also be some accompanying pragmatic information, or PI. As mentioned in §2.4 and §4.2, this will be assumed to be a binary feature styled [+/-Fam], which applies to the arguments in the clause. Therefore, the meaning input for this approach looks like the following (for e.g. LY *ko*^{55} *ɳen*^{13} *iə*^{55} *to*^{44} *tsau*^{55} *ti*^{55} *xə*^{42} *sʅ*^{12}):

[CL-N]_V_[CL-N]
PI: [-Fam], [+Fam]

The input can include any of the possible classifier-noun combinations (except Changsha, as the next section will show), with PI feature values varying based on discourse information independent of the output.

The relationship between familiarity-novelty and definiteness was established in §2.4; essentially, a positive value in one strongly prefers a positive value in the other and vice versa. This leads to the postulation of the following faithfulness constraint, formally tying the two together:

(C4) FAITHP: output definiteness should correspond to input PI

The function of this constraint is to link pragmatic considerations of familiarity with the semantic notion of definiteness. It is justified based on the existence of dialects (e.g. LY) which allow [CL-N] to have any referential interpretation; it is therefore hypothesized that discourse information must bear the burden of determining reference.

In §2.2 and §4.2, it was pointed out that Wang's (2015) typology demonstrates the markedness of preverbal [-Def] and postverbal [+Def] relative to their counterparts. Additionally, NPs tend to be limited in their referential interpretation with an accompanying bare-classifier, which harks back to the original claim that the function of bare-classifier phrases is restricting the availability of referential interpretations in favor of the unmarked interpretation (see XX, LD). There are thus three factors that must be accounted for: position, reference, and presence of a classifier. Parts of Aissen's (2003) approach to Harmonic Alignment, wherein constraints are built through alignment of markedness hierarchies, seems to provide the most straightforward method for accounting for these facts. If we have rankings postverbal [-Def] > postverbal [+Def] and preverbal [+Def] > preverbal [-Def], we can posit the following two markedness constraints[1]:

(C5) *Pos/Def ^ CL: no postverbal [+Def] bare-classifiers

(C6) *Pre/Indef ^ CL: no preverbal [-Def] bare-classifiers

The format of these constraints follows the collocation of the markedness hierarchies in §4.1 with the bare-classifier construction; '^' should be read as 'and' or 'with'. These two are mirror images of each other, in that the first disallows indefinite readings preverbally and the second disallows

[1]A reviewer points out that the postulation of these constraints is descriptive rather than explanatory. The explanatory element here is taken to be the claim that an increase in structure equates a restriction on the possible interpretations of the NP, in the form of a classifier. A more general way to formulate these would be "no postverbal definites" and "no preverbal indefinites", which could be motivated by the data. However, it seems to be the case that this does not hold as clearly with other nominal phrases, which are treated differently, which is why the more specific formulation is preferred.

definite readings postverbally. The effects of these constraints are most clearly observed in XX and LD, which are mirror images of each other in requiring [CL-N] to be limited in interpretation by position. Following Aissen's argumentation, constraints such as *Pre/Def ^ CL and *Pos/Indef ^ CL are logical equivalents to the above; however, it is unclear what their effects might be, or whether they exist at all, given that what they penalize is the preferred referential setting in Sinitic (and cross-linguistically, for that matter; Lambrecht 1994). However, it would not be such a stretch of the imagination to conceive of a classifier language which might have the opposite setting.

These constraints serve to impose a licensing restriction on marked reference value and positioning combinations. There is a large body of work on Sinitic which holds that the unmarked referential setting is S/Topic/Preposed Object = Def and O/Focus= Indef (Chao 1968: 76; Li & Thompson 1981; Xu 1997; Goddard 2005: 38; Li & Bisang 2012; Li 2013: 119).

The candidate set includes all combinations of definite and indefinite interpretations, as follows:

(16) Candidates:
a. [+Def]_V_[+Def]
b. [-Def]_V_[+Def]
c. [+Def]_V_[-Def]
d. [-Def]_V_[-Def]

Candidate (c) will be recognized as the maximally unmarked option, with a preverbal definite and a postverbal indefinite, while candidate (b) will be recognized as the maximally marked option with the opposite configuration. Candidates (a) and (d) are equal in terms of markedness.
We now have enough theoretical machinery to account for the XX, LD, and LY dialects. These dialects differ fundamentally from CS in allowing [CL-N] to occur in all positions. Therefore, for reasons of space, illustrative tables are provided only for strings of the {[CL-N]_V_[CL-N]; [-Fam], [+Fam]} type, which is where these dialects differ from each other. Beginning with XX, consider the following ranking, with no licensing of preverbal indefinites, requiring *Pre/Indef ^ CL to be undominated:

XX**:** *Pre/Indef ^ CL >> FAITHP >> *Pos/Def ^ CL

This is illustrated in the following table:

Table 3. Unidirectional approach to XX perceptual output

[CL-N]_V_[CL-N] PI: [-Fam], [+Fam]	*Pre/Indef ^ CL	FAITHP	*Pos/Def ^ CL
→ a. [+Def]_V_[+Def]		*	*
b. [-Def]_V_[+Def]	*!		*
c. [+Def]_V_[-Def]		**!	
d. [-Def]_V_[-Def]	*!	*	

Candidates (b-d) are both ruled out due to the highly ranked *Pre/Indef ^ CL. Meanwhile, FAITHP must militate against an interpretation at odds with the pragmatic information, which determines candidate (a) as the winning candidate, since (c) does not agree with PI. Changing of the PI features would allow the following candidates to surface: [+Fam, -Fam]; [+Fam, +Fam] = (c); [-Fam, +Fam]; [-Fam, -Fam] = (a).

The LD dialect is a mirror image of the one seen above, in that it disallows postverbal definites; therefore, *Pos/Def ^ CL must be ranked above all other constraints, as seen in the following ranking:

LD: *Pos/Def ^ CL >> FAITHP >> *Pre/Indef ^ CL

Consider the following table:

Table 4. Unidirectional approach to LD perceptual output

[CL-N]_V_[CL-N] PI: [-Fam], [+Fam]	*Pos/Def ^ CL	FAITHP	*Pre/Indef ^ CL
a. [+Def]_V_[+Def]	*!	*	
b. [-Def]_V_[+Def]	*!		*
c. [+Def]_V_[-Def]		**!	
→ d. [-Def]_V_[-Def]		*	*

As with XX, FAITHP must militate against an interpretation at odds with the pragmatic information, which determines candidate (d) as the winning candidate. Changing of the PI features would allow the following candidates to surface: [+Fam, -Fam]; [+Fam, +Fam] = (c); [-Fam, +Fam]; [-Fam, -Fam] = (d).

Both XX and LD demonstrate the effect of pragmatic information in choosing candidates when a bare classifier cannot restrict the referential interpretation due to a lower ranking. LY takes this one step further, since its classifiers impose no restrictions on NP reference (Chen 1999 and Wang 2015); therefore, the burden of determining reference falls entirely on discourse or pragmatic considerations of familiarity-novelty. The corresponding ranking would require FAITHP to be undominated by any other constraint, while *Pre/Indef ^ CL and Pos/Def ^ CL are unranked with respect to each other; consider:

LY: FAITHP >> *Pre/Indef ^ CL, *Pos/Def ^ CL

This is illustrated in table 5:

Table 5. Unidirectional approach to LY perceptual output

[CL-N]_V_[CL-N] PI: [-Fam], [+Fam]	FAITHP	*Pre/Indef ^ CL	*Pos/Def ^ CL
a. [+Def]_V_[+Def]	*!		*
→ b. [-Def]_V_[+Def]		*	*
c. [+Def]_V_[-Def]	**!		
d. [-Def]_V_[-Def]	*!	*	

Since this dialect allows all possible iterations of [CL-N] and definiteness, the deciding factor in determining reference value is checking of the value for PI; i.e. it is hypothesized that pragmatics is the sole arbiter when it comes to determining reference value. Any of the other candidates can surface if the PI features are changed to correspond to appropriate definiteness values, regardless of the presence or lack of a classifier.

Finally, CS is somewhat of an outlier; while it has an LD-like restriction on the interpretation of postverbal [CL-N], it has no restriction on the reference value of the preverbal, since it has an outright ban (see §3.1 and §4.2) on the appearance of [CL-N] in this position. For this reason, its input must be of the [N]_V_[CL-N] or [N]_V_[N] type, since it is predicted that [CL-N]_V_[CL-N] or [CL-N]_V_[N] inputs will not be encountered by CS speakers. While it is impossible to see the effects of *Pre/Indef ^ CL, it is believed that at the level of comprehension CS functions much like LD. Consider the following ranking:

CS: *Pos/Def ^ CL >> FAITHP >> *Pre/Indef ^ CL

This is illustrated in the following table:

Table 6. Unidirectional approach to CS perceptual output

[N]_V_[CL-N] PI: [-Fam], [+Fam]	*Pos/Def ^ CL	FAITHP	*Pre/Indef ^ CL
a. [+Def]_V_[+Def]	*!	*	
b. [-Def]_V_[+Def]	*!		
c. [+Def]_V_[-Def]		**!	
→ d. [-Def]_V_[-Def]		*	

Two aspects are unique to CS: the input is limited to preverbal [N] by the winning candidate of form optimization, indicating that there is a direct relationship between it and meaning optimization (see §4.2 and §4.4), and consequently the *Pre/Indef ^ CL constraint sees no effect (for which reason it is believed to be ranked lowest). Candidate (d) is predicted to win based on its minimal violation of FAITHP, although as with LD variation of the features may lead to candidate (c) surfacing.

While the facts about meaning optimization in the three preceding dialects could be accounted for solely with this unidirectional account, the CS data is additionally restricted at the production stage, and therefore a full account of bare-classifiers in these dialects requires reference to both the semantic-pragmatic and syntactic levels (see §4.1). We therefore posit that in CS the production stage operates first, providing the output for the perception stage. This is unproblematic if we assume sequential optimization of form and meaning (following Kuhn 2003), but adds an additional layer of complication if we assume simultaneous optimization as found in most bidirectional approaches (see Blutner 2000; Lestrade, Van Bergen, and De Swart 2016). A suitable outline of this second type

of account, however, is outside the scope of the present piece; this is discussed in §5.

5. DISCUSSION AND CONCLUSIONS

The present study represents an attempt at formalizing the variation of Xiang bare-classifier constructions from the perspective of Optimality Theory, following typological-comparative work found in Wang (2015). Several key assumptions about the function of classifiers were adopted, including the following:

1. Bare-classifiers are not yet fully grammaticalized as markers of reference, as evidenced by an apparent lack of consistency in their interpretation as definite or indefinite;
2. Bare-classifiers only unambiguously mark individuation of nouns, just as in their use following numerals, quantifiers, and demonstratives; and
3. The function of classifiers as regards reference is one of limiting the amount of possible referential interpretations, with a preference for the unmarked interpretation given a certain position relative to the verb.

Each of the Xiang dialects under investigation involved a step on the spectrum from most to least freedom in the interpretation of classifiers, from Lianyuan, with all possible referential interpretations by position, to Changsha, which limits their appearance and interpretation to only postverbal indefinite. In between these two are Xiangxiang and Loudi, which allow all interpretations but one: preverbal indefinites for the former, postverbal definites for the latter. This was formalized in terms of constraints mediating syntactic positioning of the [CL-N] construction and constraints requiring the marking of individuation in a unidirectional approach to form optimization, and in terms of constraints mediating semantic interpretation by position in a unidirectional approach to meaning optimization. While the latter could account for the facts about

LY, XX, and LD, the combination of both unidirectional approaches is necessary to account for the facts about CS.

In answering the question of why certain positions interact with bare-classifier reference in these dialects, it seems to be the case that the addition of structure limits the available interpretations for bare-NPs in Sinitic. For instance, a bare-NP in the LD dialect can be thought of as being indeterminate for the following features in postverbal position: [~Ind], [~Def], etc... The addition of a classifier will restrict the interpretation to [+Ind], [-Def]; but, preverbally, the interpretation would be [+Ind], [~Def]. Why? The answer adopted here is that classifiers *mark* individuation but only *interact* with definiteness through a restriction to an unmarked setting. Dialects differ in terms of how or whether classifiers restrict interpretation; in LY, classifiers *only* mark individuation and do not interact with definiteness, while in XX and LD preverbal indefinites and postverbal definites are particularly excluded, respectively. Why a certain positional reference is marked in particular is considered essentially arbitrary and dialect-specific and is treated as a dimension along which dialects vary.

5.1 Extensions to Simultaneous Bidirectional OT

The sequential unidirectional approach as it now stands could be expanded on to include a fully simultaneous bidirectional one, given the fundamental differences in the treatment of [CL-N] in CS as opposed to the other dialects. Unlike in XX, LD, and LY, which have no syntactic restrictions on [CL-N], CS interpretation of [CL-N] is heavily dependent on syntax, in that there is an arbitrary ban on preverbal [CL-N] which constrains the possible inputs. For this reason, a proposed final step in accounting for this difference is the formal combination of the two unidirectional approaches into a single bidirectional one; simultaneously optimizing form for meaning, and meaning for form. Following original work in Blutner (2000), and more recent work in Lestrade, Van Bergen, and De Swart (2016) dealing with the interface of syntax-semantics, an asymmetrical bidirectional approach to the optimization of production and perception could be adopted for the present analysis. In contrast to traditional unidirectional approaches to syntax and semantics (c.f.

Grimshaw 1997; Aissen 2003; de Hoop and Lamers 2006), this approach matches form-meaning pairs, thus taking into account both speaker and listener effects in output optimization, both of which are considered essential for the production of syntactically and semantically optimal constructions.

Naturally, this approach requires more theoretical machinery to formalize the feedback loop of production and perception. The current approach displays greater formal simplicity but encounters several issues in accounting for the possible optionality of the data, which might be solved through recourse to bidirectionality. However, there are several reasons why this is impractical. For one, it is not clear that definiteness is *directly* recoverable from differing iterations of [CL-N] and [N] in the same way that predicate semantics is. In approaches such as that in Lestrade, Van Bergen, and De Swart (2016), who are dealing with differential object marking and transitive clauses, the input for production is the same as the output for interpretation, and vice versa. It is clear how propositional inputs like BUY(x, y) are directly retrievable from outputs like 'x bought y' or 'y bought x'. However, as mentioned above, 'definiteness', as a 'semantic property' of nouns (Frawley 1992: 68), straddles both semantics (by assigning properties to referents, e.g. uniqueness, familiarity) and pragmatics (by selecting or identifying discourse referents) (Frawley 1992: 71-72), and has no direct relationship to argument semantics (or, in this approach, to output candidates at the production level). Therefore, the feedback loop necessary for bidirectional optimization does not seem to hold in this case, and further work would be necessary to successfully implement this approach.

5.2 Extensions to Other Reference Marking Strategies[2]

In developing the current approach to bare-classifier syntax it is worth noting how other methods of reference marking in Sinitic can be integrated into this account. These can include use of verbal auxiliaries, movement and topicalization, phrase-types which may mark a certain referential value, among others. §3.5 briefly mentioned the introduction

[2]This section was inspired by comments from reviewers on other reference marking strategies and nominal phrases in Sinitic.

of indefinite NPs with 'have'-type verbal auxiliaries (referred to here as the 'Presentational Construction', following Wiedenhof 2015), and similar processes include the definiteness-marking *ba*-construction (referred to here as the 'Disposal Construction', following Wu 2005). While data on how these Xiang dialects treat the presentational construction is not currently available, it may be illustrative to use Mandarin as a baseline for comparison. Consider the following two constructions from MD:

(17) a. *yǒu táozi gǔn xiàlái*
have peach roll down
'*A* peach rolled down'
b. *yǒu ge táozi gǔn xiàlái*
have CL peach roll down
'*A* peach rolled down'

The construction in (a) provides the NP *táozi* 'peach' with an obligatory indefinite interpretation, by preceding it with the verbal element *yǒu* 'have', while the construction in (b) shows the same process co-occurring with the bare-classifier construction. As seen previously, if the [CL+N] construction were to occur without the verbal element, it would be ungrammatical.

In addition, consider the following disposal constructions from CS (a) and LD (b), from Wu (2005: 188, 206):

(18) a. $\underline{pa^{41}}$ $tɕ^{h}yan^{33}fu$ ta^{41} $k^{h}ai^{33}$
DISP window open COMP
'Open the window (please)'
b. $ŋ^{42}$ $\underline{nõ^{44}tɤ^{2}}$ ti^{5} $tɕĩ^{13}$ $tɕ^{h}ĩœ^{35}$ $ka^{5}li^{5}$
1.sg DISP CL money lose ASP
'I lost the money'

Here the preposing of the NP with the disposal marker gives it a definite interpretation. Interestingly in (18b), just as with (17b) above, we can see

that LD allows the disposal construction and the bare-classifier construction to co-occur, which lends further credence to our claim that classifiers are not markers of definiteness. If they were, there would be no need to add either a disposal marker or a verbal auxiliary to supplement the classifier.

The crucial similarity between these two processes and the bare-classifier construction is that all of them are dependent on positioning relative to the verb, rather than simply the presence of a marker, in order to gain their reference value. Consider that indefinites in the presentational construction *must follow* a verbal element in order to gain their referential value, while similarly the disposal 'ba'-construction introduces a definite NP which *must precede* the main verb. Few, if any, would claim that *you* 'have' in MD is an 'indefinite determiner', or that *ba* is a 'definite determiner', despite their association with these referential values; it is the syntactic positioning of the NP which provides its reference. Essentially, the same can be said about bare-classifier constructions.

In developing a unified account of these three phenomena, one could posit more general constraints which penalize non-canonical reference values in certain positions, which could motivate the use of verbal auxiliaries or movement with *ba* in order to exploit position relative to the verb for reference marking purposes. For instance, one could see how the following two constraints:

(C5) *Pos/Def ^ CL

(C6) *Pre/Indef ^ CL

…could be reformulated as:

*Non-Canonical: penalize any appearance of non-canonical reference

This sort of expansion is tempting to consider, given the similarity between these three referential constructions. However, given the lack of decisive data on how presentational constructions occur in Xiang, the issue is left open here. Further research on Xiang and position-dependent reference marking strategies is necessary.

Other types of nominal phrases, such as NumPs or DPs, are also presumably not restricted in the same way as ClPs, but a lack of information on how these dialects treat these items has made it hard to make sizable claims. Evidence from Mandarin can illustrate the

differences, however: both NumPs and DPs can occur in any position relative to the verb, and NumPs can have any referential value dependent on position, topicalization, etc. (see discussion in Li 2013). DPs are naturally exclusively definite (depending on one's theoretical background, here following Lyons 1999), owing to the marking of [+Def] reference on the NP by items such as *zhe* and *na*. Clearly there are some fundamental differences between these three types of phrases; but, again, it is hard to say how exactly this applies to other nominal phrases in the dialects in question.

5.3 Bare-Classifier Typology

The OT account outlined in §4 allows for the development of a factorial typology of referential classifier constructions, a considerable strength when compared to other theories of grammatical variation (Kager 1999; Kuhn 2003). A typology of this type, adopted for Wang's (2015) typology, may be extended further to encompass other Sinitic varieties outside Xiang. Given that the two constraint sets (form-optimizing and meaning-optimizing) do not directly interact, we can predict that the number of possible rankings will be F = 6, M = 6 (where n = number of constraints, and n! = number of possible rankings; Kager 1999), consider the following table:

Table 7: Possible constraint rankings, by optimization domain.

Form	Meaning
1) *POS-CL >> *PRE-CL >> FAITHIND 2) *PRE-CL >> *POS-CL >> FAITHIND 3) *PRE-CL >> FAITHIND >> *POS-CL 4) FAITHIND >> *POS-CL >> *PRE-CL 5) FAITHIND >> *PRE-CL >> *POS-CL 6) *POS-CL >> FAITHIND >> *PRE-CL	1) FAITHP >> *Pos/Def ^ CL >> *Pre/Indef ^ CL 2) *Pos/Def ^ CL >> FAITHP >> *Pre/Indef ^ CL 3) *Pos/Def ^ CL >> *Pre/Indef ^ CL >> FAITHP 4) *Pre/Indef ^ CL >> *Pos/Def ^ CL >> FAITHP 5) *Pre/Indef ^ CL >> FAITHP >> *Pos/Def ^ CL 6) FAITHP >> *Pre/Indef ^ CL >> *Pos/Def ^ CL

By combining the constraint sets for all dialects, we can see that the below rankings correspond to Types I, II, III, and VII in the typology:

Type I: FAITHP, FAITHIND >> *Pos/Def ^ CL, *Pre/Indef ^ CL >> *PRE-CL, *POS-CL

Type II: *Pos/Def ^ CL, FAITHIND >> FAITHP >> *Pre/Indef ^ CL >> *PRE-CL, *POS-CL

Type III: *Pre/Indef ^ CL, FAITHIND >> FAITHP >> *Pos/Def ^ CL >> *PRE-CL, *POS-CL

Type VII: *PRE-CL, *Pos/Def ^ CL >> FAITHIND, FAITHP >> *POS-CL, *Pre/Indef ^ CL

Other variations on these rankings also produce the other three types. For instance, Types IV (only unmarked postverbal definite and preverbal

definite, V (only postverbal definite and indefinite), and VI (no [CL-N]) can be derived with the following rankings:

Type IV: FAITHIND, *Pos/Def ^ CL, *Pre/Indef ^ CL >> FAITHP >> *PRE-CL, *POS-CL

Type V: *PRE-CL >> FAITHIND, FAITHP >> *Pos/Def ^ CL, *Pre/Indef ^ CL >> *POS-CL

Type VI: *POS-CL, *PRE-CL >> FAITHIND, FAITHP, *Pos/Def ^ CL, *Pre/Indef ^ CL

The remaining rankings predict other dialect types that do not exist or have not yet been determined to exist. In addition, certain rankings end up having the same output (for instance any ranking where *POS-CL and *PRE-CL outrank FAITHIND). Further research could expand upon the possible/impossible alternatives, and further generalization of the constraints involved to different contexts.

5.4 Limitations and Conclusions

The research as it is presented here relies heavily on the typology in Wang (2015) to determine grammaticality/ungrammaticality; the actual picture in the Xiang dialects may be considerably more variable or complex. In order to make claims with more certainty, a greater degree of fieldwork is required, focusing specifically on the interpretation of classifier modified NPs in different positions. This is particularly true when dealing with shades of meaning as open to reinterpretation as reference.

In terms of the theoretical approach, the constraints as they are posited rely solely on Sinitic data, which provides grounds for their existence, but limits their scope of application. It is quite possible that if the scope of the paper were extended to classifier languages outside of Sinitic (such as classic work of the type found in Bisang 1999), then more general trends could be encoded into the constraints under consideration, rather than relying on constraints solely focused on classifiers. As has been pointed

out to the author (Cathryn Donohue p.c.), it seems likely that there are some more basic features underlying classifiers that may be limiting their distribution in particular syntactic contexts, rather than just their status as markers of individuation.

On a similar note, a reviewer points out that the constraints as they are formulated are primarily descriptive of linguistic facts, rather than explanatory. This is particularly true of the two constraints which limit interpretation by position, the *Pos/Def ^ CL and *Pre/Indef ^ CL constraints. It should be noted here that the purpose of these constraints is to formalize the typological variation, rather than explain it, since the explanation for the facts is held to stem from the function of classifiers. A possible alteration to these constraints which would make them more generalizable is provided in the note in §4.2, but the concern is that while it may be able to apply to more phenomena, it may not take account of the unique ways bare-classifier constructions are treated.

Related to this, the constraints pinpoint a very specific aspect of Sinitic grammar (bare-classifiers), rather than a general trend (such as NP reference marking). Of course, the relative prominence of preverbal material over postverbal material is a crosslinguistic trend, but the fact that Sinitic varieties seem to restrict one position over the other in a seemingly arbitrary way presents an issue. For instance, the 'mirror image' dialects, Xiangxiang and Loudi, prioritize restrictions of reference either in preverbal or postverbal position, respectively, which must be attributed to linguistic variation within Sinitic. This is problematic if we assume the traditional OT view that constraints are cross-linguistically universal (e.g. Prince & Smolensky 1993; Kager 1999); it would then assume that all languages have some form or other of these constraints somewhere in their grammar. More recent theories of OT acquisition, however, have proposed that specific constraints are created and learned as the language is acquired, without the need for recourse to universal constraints (e.g. Doyle et al. 2014). Regardless, it is held that the constraints are grounded in (Sinitic) typological trends if nothing else, which must be accounted for in any case.

The overall purpose of the work has been to capture the observed variation across Xiang in a formal theoretical framework, in such a way that the factors believed to be underlying their distribution are incorporated effectively. At its heart, it is an attempt to explore the

grammatical diversity of Sinitic, focusing on an understudied subgrouping, while simultaneously expanding on our understanding of classifier usage generally.

REFERENCES

Aikhenvald, A. 2000. *Classifiers: A Typology of Noun Categorization Devices*. Oxford University Press.

Aissen, J. 2003. Differential Object Marking: Iconicity vs. Economy. *Natural Language & Linguistic Theory, 21*(3), 435-483.

Bisang, W. 1999. Classifiers in east and southeast Asian languages: Counting and beyond. *Numeral Types and Changes Worldwide*. Berlin: Mouton de Gruyter. 113-185.

Bisang, W. & Wu, Y. 2017. Numeral classifiers in east Asia. *Linguistics 55*(2), 257-264.

Blutner, R. 2000. Some aspects of optimality in natural language interpretation. *Journal of Semantics* 17(3): 189-216.

Bybee, J., Pagliuca, W., & Perkins, R. 1990. On the asymmetries in the affixation of grammatical material. In Croft, Denning, and Kemmer (eds.), *Studies in Typology and Diachrony*, 1-42. Amsterdam: John Benjamins.

Chafe, W. 1987. Cognitive constraints on information flow. *Coherence and Grounding in Discourse*, 11, 21-51.

Chao, Y. R. 1968. *A Grammar of Spoken Chinese*. Berkeley: University of California Press.

Chappell, H. M. 2001. Synchrony and diachrony of Sinitic languages: A brief history of Chinese dialects. In Chappell (ed.), *Sinitic Grammar: Synchronic and Diachronic Perspectives*, 3-29. Oxford University Press.

Chappell, H. M. 2015. Linguistic areas in China for differential object marking, passive, and comparative constructions.' In Chappell (ed.), *Diversity in Sinitic Languages*, 13-53. Oxford University Press.

Chen, C. 1958. Minnanhua he Putonghua changyong liangci de bijiao [A comparison of frequently used classifiers in Southern-Min and Standard Mandarin]. In *Zhongguo Yuwen* 1958/12, 591-593.

Chen, H. 1999. *Lianyuan fangyan yanjiu* [Research on the Lianyuan Dialect]. Changsha: Hunan Education Press.

Cheng, L. L. S. & Sybesma, R. 1999. Bare and not-so bare nouns and the structure of NP'. *Linguistic Inquiry* 30(4): 509-542.

Cheng, L. L. S. & Sybesma, R. 2005. Classifiers in four varieties of Chinese. In *Handbook of Comparative Syntax*, 259-292.

Cheng, L. L. S. & Sybesma, R. 2014. The syntactic structure of noun phrases. In Huang, Li, and Simpson (eds.), *Handbook of Chinese Linguistics*, 26-49. Chichester, UK: Wiley Blackwell.

Chierchia, G. 1998. Plurality of mass nouns and the notion of "semantic parameter". In Rothstein (ed.), *Event and Grammar*, 53-103. Dordrecht: Kluwer.

Choi, H.Y. 2001. Binding and discourse prominence: Reconstruction in "focus" scrambling. In Legendre, Grimshaw, and Vikner (eds.), *Optimality-Theoretic Syntax*. Cambridge, MA: MIT Press.

Del Gobbo, F. 2014. Classifiers. In Huang, Li, and Simpson (eds.), *Handbook of Chinese Linguistics*, 26-49. Chichester, UK: Wiley Blackwell.

Doyle, G., Bicknell, K. & Levy, R. 2014. Nonparametric learning of phonological constraints in optimality theory. *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*. Stroudsburg, PA: Association for Computational Linguistics.

Erbaugh, M. S. 1986. Taking stock: The development of Chinese noun classifiers historically and in young children. In Craig (ed.), *Noun Classes and Categorization*, 399-436. John Benjamins.

Erbaugh, M. S. 2013. Classifier choices in discourse across the seven main Chinese dialects. In *Increased Empiricism: Recent Advances in Chinese Linguistics*, 101-126.

Frawley, W. 1992. *Linguistic Semantics*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Goddard, C. 2005. *The Languages of East and Southeast Asia: An Introduction*. Oxford University Press.

Grimshaw, J. 1997. Projection, heads, and optimality. *Linguistic Inquiry* 28: 373-422.

He, W. 2006. Xiangxiang (Hutian) huà zhōng "zhī + míng" jiégòu yánjiū [Study on the "zhi + Noun" construction in Xiangxiang (Hutian)]. Retrieved from: http://gb.global.cnki.net/index/. *Zhongguo Zhiwang*. Master's Dissertation.

Her, O. S. & Hsieh, C. T. 2010. On the semantic distinction between classifiers and measure words in Chinese. *Language and Linguistics*, *11*(3), 527-551.

de Hoop, H. & Lamers, M. 2006. Incremental distinguishability of subject and object. *Case, Valency and Transitivity*, 269-287.

Hsieh, M. L. 2008. *The Internal Structure of Noun Phrases in Chinese*. Taipei: Crane.

Kager, R. 1999. *Optimality Theory*. Cambridge University Press.

Kuhn, J. 2003. *Optimality Theoretic Syntax: A Declarative Approach.* Stanford, CA: CSLI.

Lambrecht, K. 1994. *Information Structure and Sentence Form*. Cambridge University Press.

LaPolla, R. J. 2003. Why languages differ: Variation in the conventionalization of constraints on inference. In *Language Variation: Papers on variation and change in the Sinosphere and in the Indosphere in honour of James A. Matisoff*, 113-144.

Legendre, G., Wilson, C., Smolensky, P., Homer, K. & Raymond, W. 1995. Optimality and wh-extraction. In Beckman, Dickey, & Urbanczyk (eds.), *University of Massachusetts Occasional Papers 18: Papers in Optimality Theory*, 607-636. Amherst: University of Massachusetts Press.

Legendre, G. 2001. An introduction to optimality theory in syntax. In Legendre, Grimshaw, & Vikner (eds.), *Optimality-Theoretic Syntax*. Cambridge, MA: MIT Press.

Legendre, G., Putnam, M., De Swart, H. & Zaroukian, E. 2016. Introduction. In Legendre, Putnam, De Swart, & Zaroukian (eds.), *Optimality-Theoretic Syntax, Semantics, and Pragmatics: From Uni- to Bidirectional Optimization*, 1-33. Oxford University Press.

Lestrade, S., van Bergen, G. & de Swart P. 2016. On the origin of constraints. In Legendre, Putnam, De Swart, & Zaroukian (eds.), *Optimality-Theoretic Syntax, Semantics, and Pragmatics: From Uni- to Bidirectional Optimization*, 179-99. Oxford University Press.

Li, X. 2013. *Numeral Classifiers in Chinese: The Syntax-Semantics Interface*. Berlin: De Gruyter Mouton.

Li, X. & Bisang, W. 2012. Classifiers in Sinitic languages: From individuation to definiteness marking. *Lingua* 122, 335-355.

Li, C. & Thompson, S. 1981. *A Functional Reference Grammar of Mandarin Chinese*. Berkeley: University of California Press.

Liu, Lihua. 2001. *Loudi fangyan yanjiu* [Research on the Loudi Dialect]. Changsha: Central South University Press.

Longobardi, G. 1994. 'Reference and Proper Names'. *Linguistic Inquiry* 25(4).

Lü, S. 1944. *Ge* zi de yingyong fanwei, fulun danweici qian *yi* zi de tuoluo [On the usages of GE and the dropping of the numeral YI before classifiers]. *Xihua Daxue zhongguo wenhua huikan*. Jinlin: Qilu.

Lyons, C. 1999. *Definiteness*. Cambridge University Press.

Matthews, S. & Yip, V. 2013. *Cantonese: A Comprehensive Grammar*. London: Routledge.

Norman, J. 1988. *Chinese*. Cambridge University Press.

Packard, J. 2000. *The Morphology of Chinese: A Linguistic and Cognitive Approach.* Cambridge University Press.

Prince, A. & Smolensky, P. 1993; 2004. *Optimality Theory: Constraint Interaction in Generative Grammar*. Rutgers Optimality Archive 537, roa.rutgers.edu.

Ramsey, S. R. 1989. *The Languages of China*. Princeton University Press.

Simons, G. & C. Fennig. (eds). 2018. Chinese, Xiang. *Ethnologue: Languages of the World, Twenty-first edition*. Dallas: SIL International. http://www.ethnologue.com.

Simpson, A. 2005. Classifiers and DP structure in southeast Asia. In Cinque & Kayne (eds.), *The Oxford Handbook of Comparative Syntax*, 806-838. Oxford University Press.

Sio J. U. 2007. *Modification and Reference in the Chinese Nominal*. Utrecht: Netherlands

Graduate School of Linguistics. Doctoral Dissertation.
Wang, J. 2015. Bare classifier phrases in Sinitic languages: A typological perspective. In Chappell (ed.), *Diversity in Sinitic Languages*, 110-133. Oxford University Press.
Wiedenhof, J. 2015. *A Grammar of Mandarin*. Amsterdam: John Benjamins.
Wu, Y. 1999. The development of aspectual systems in the Chinese Xiang dialects. *Collection des Cahiers de Linguistique Asie Orientale, 4*. Paris: Ecole des Hautes Etudes en Sciences Sociales.
Wu, Y. 2005. *A Synchronic and Diachronic Study of the Grammar of the Chinese Xiang Dialects*. Berlin: Mouton de Gruyter.
Xu, L. 1997. Limitations on subjecthood of numerically quantified noun phrases: A pragmatic approach. In Xu (ed.), *The Referential Properties of Chinese Noun Phrases*, 25-45. Paris: Ecole des Hautes Etudes en Sciences Sociales.
Zhou, C. 1991. *Minnanhua yu Putonghua* [Southern-Min and Standard Mandarin]. Yuwen: Peking.

Robert Marcelo Sevilla
Department of Linguistics
The University of Hong Kong
Pokfulam Road, Hong Kong
u3545575@connect.hku.hk

湘語的量名結構
從類型學到形式化

Robert Marcelo Sevilla
香港大學

漢語量名結構在名詞前面只有量詞而沒跟隨數詞或指示代詞的時候，展演一套句法與語意的複雜相互作用。此結構與指稱狀況標記相關聯，表面上對於能用在何處及何種演繹只存在任意的限制。本研究介紹湘語方言中量名結構的變化，從而推展湘語量詞結構的類型學 (套用 Wang 2015 的類型學研究)，以納入制約優選理論的形式化框架。從所得句法與語意制約推論，量名結構並不標記指稱狀況，而是限制名詞的指稱詮釋可能性，同時只會直接標記度量單位。本文通過四個湘語方言作為代表：長沙、湘鄉、婁底和漣源，展示量名結構從全部語境通用，到只能用於動詞後的無定指稱的廣闊多樣性。形式化論述兼用句法-語意雙向優選理論的制約闡釋有效地捕捉到類型學趨勢。

關鍵字：量名結構、湘語、類型學、句法-語意、優選理論