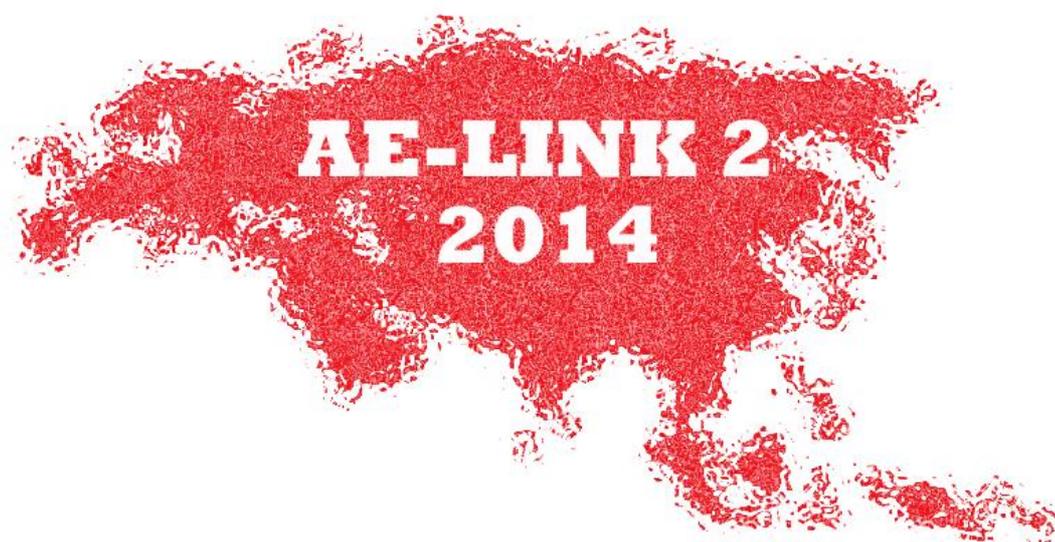


The Second Asian and European Linguistic Conference Proceedings

Edited by
Alison Biggs, Man Li, Aiqing Wang and Cong Zhang



NNWPL

Volume 21, Issue 1
2015

NEWCASTLE AND NORTHUMBRIA WORKING PAPERS IN LINGUISTICS

ISSN 2041-1057

NEWCASTLE UNIVERSITY LIBRARY

ISBN 978-0-7017-0251-9

EDITORIAL BOARD

Alison Biggs
Man Li
Aiqing Wang
Cong Zhang

PUBLISHED BY

Centre for Research in Linguistics and Language Sciences (CRiLLS)
Newcastle University
AND
Newcastle University Library

FOR INQUIRIES

CRiLLS Room 3.01 & 3.02, Level 3,
Old Library Building
Newcastle University
Newcastle upon Tyne
NE1 7RU
United Kingdom
Telephone: 0191 222 8790
Fax: 0191 222 6592
E-mail: crills@ncl.ac.uk

TABLE OF CONTENTS

Preface	v
The Editors	
The Word Order Of The Directional Serial Verb Construction In Mandarin Chinese Zhishuang Chen	1
Two Types Of States In Korean: Experimental L1 Evidence Jiyoung Choi	17
Licensing Definiteness: A Spanning Account Of Noun Phrase Interpretation In Mandarin Chinese And Cantonese David Hall	31
Children’s Interpretation Of Japanese Particles In Complex Sentences Keiko Hata	47
Generic Pronouns And Phi-Features: Evidence From Thai Anders Holmberg and On-Usa Phimsawat	55
Affect In Chinese And Korean Spoken Narratives Vivian Lee and Lu Lu	72
Word Formation In Phase Theory Makiko Mukai	84
From “Hen” To Adjectival Modification In Mandarin Chinese Fangfang Niu	96
Syntactic And Semantic Change In Chinese Alain Peyraube	112
Facts: The Interplay Between The Matrix Predicate And Its Clausal Complement Ji Young Shim and Tabea Ihsane	130
<i>Wh</i> -Scope Marking In Syrian Arabic: An Indirect Dependency Mais Sulaiman	145
The Intervention Effect Of Negation On <i>Wh</i> -Adverbials In Late Archaic Chinese Aiqing Wang	160

PREFACE: NOTE FROM THE EDITORS

ALISON BIGGS
(University of Oxford)

MAN LI
(Newcastle University)

AIQING WANG
(University of York)

CONG ZHANG
(University of Oxford)

The Second Asian and European Linguistic Conference (AE-Link 2) took place at Newcastle University on 5-6th December 2014. The purpose of the conference was to provide a forum for the presentation and discussion of the best current research on the languages of Asia, and to promote interdisciplinary collaboration between linguists from Europe and Asia. This Special Issue collects together a selection of the papers presented at the conference.

The editors would like to take this opportunity to express thanks to the participants for making this such a successful event. We would also like to extend thanks to the many people who contributed to the organisation of the conference, in particular the many reviewers involved in the original selection of papers. For their role as conference volunteers, we would like to thank Daniel Bell, Dongyan Chen, Harold Thampoe, Azad Maudaressi, Rebecca Musa, Qi Chen, Danhui Li, Ruoyao Zhang, Tianche Li, Zeying Jin, Ziyu Zhou. Finally we would like to thank our sponsors: Quality in Learning and Teaching (QuILT) at Newcastle University, The Linguistics Association of Great Britain (LAGB), the Centre for Research in Linguistics and Language Sciences (CRiLLS) at Newcastle University, and the Chinese Students & Scholars Association – Newcastle (CSSA). The conference would not have been possible without this support.

The Editors

THE WORD ORDER OF THE DIRECTIONAL SERIAL VERB CONSTRUCTION IN MANDARIN CHINESE

ZHISHUANG CHEN
(University of York)

Abstract

This paper investigates the word order alternation demonstrated by the directional serial verb construction (DSVC) in Mandarin Chinese. It is observed that the word orders differ in some syntactic and semantic properties, hence are not freely interchangeable. The paper proposes two types of merge structures underlying the different word orders, specifically, a resultative structure and a double VP structure. All the syntactic and semantic divergences observed in the word orders can be derived with this approach.

1. Introduction

Mandarin Chinese is a language rich in serial verb constructions. The focus of this paper is on a special type of them, which I term as ‘directional serial verb construction’ (DSVC). A DSVC consists of a string of verbs ‘ V_1V_2 ’ in which V_1 denotes motion/displacement while V_2 denotes the direction of V_1 . DSVCs can be divided into three types according to the direction that V_2 denotes (Li & Thompson 1981).

Type 1: V_2 denotes deictic direction. There are only two of them: *lai* ‘come/towards the speaker’ and *qu* ‘go/away from the speaker’.

- (1) Ta song lai tang le.
he send come soup SFP¹
‘He brought the soup.’

Type 2: V_2 denotes non-deictic direction. There are eight verbs in this group: *shang* ‘ascend/up’, *xia* ‘descend/down’, *jin* ‘enter/into’, *chu* ‘exit/out of’, *qi* ‘rise/up’, *hui* ‘return/back’, *guo* ‘cross/over’, *kai* ‘open/apart-away’.

- (2) Ta zou shang le louti.
he walk ascend ASP stairs
‘He walked up stairs.’

Type 3: V_2 is a directional verb cluster composed of a non-deictic and a deictic directional verb. Traditionally, the directional verb cluster is viewed as a compound. A hyphen is used to connect the morphemes within the compound.²

¹ There are two *les* in the literature, the aspect marker *le* which immediately follows a verb, and the sentence final particle *le* which occurs at the end of the sentence. However, whether the *le* following an intransitive verb at the end of the sentence is an aspect marker or a sentence final particle or both remains a puzzle. In this paper I will gloss *le* at the end of a sentence that does not end with a verb as SFP (sentence final particle), and gloss *le* directly following a verb (including verbs in the middle of the sentence and those at the end of the sentence) as ASP.

² Some readers may question the compound status of the cluster *shang lai* and argue that they are two independent verbs. I will argue in Section 4 that the directional verb cluster is indeed a compound.

- (3) Ta zou shang-lai le.
 he walk ascend-come ASP
 'He walked up (towards the speaker).'

2. The Puzzle

2.1. The Word Order Alternation

Li and Thompson (1981) and many traditional descriptive linguists have noticed that different word orders are possible when constructing a DSVC. I illustrate the word order alternation allowed for each type of DSVC below, with a word order schema in the bracket at the end of each example (V_m = motion/displacement verb, V_d = directional verb):

The type 1 DSVC allows two word orders.

- (4) a. Ta song lai le tang. ($V_m V_d O$)
 he send come ASP soup
 b. Ta song le tang lai. ($V_m O V_d$)
 he send ASP soup come
 'He brought the soup.'

The type 2 DSVC allows only one order.

- (5) a. Ta zou shang le louti. ($V_m V_d O$)
 he walk ascend ASP stairs
 b. *Ta zou le louti shang. ($V_m O V_d$)
 he walk ASP stair ascend
 'He walked up stairs.'

The type 3 DVSC allows three word orders.

- (6) a. Ta song jin-lai le tang. ($V_m V_d-V_d O$)
 he send enter-come ASP soup
 b. Ta song le tang jin-lai. ($V_m O V_d-V_d$)
 he send ASP soup enter-come
 c. Ta song jin le tang lai. ($V_m V_d O V_d$)
 he send enter ASP soup come
 'He brought the soup into here.'

What interests us is the word alternations shown by the type 1 and 3 DSVCs. The examples above are representative of the previous literature. The divergences between the word orders in (4) and (6) respectively receive no further investigation. The examples seem to suggest that the order alternations are equivalent. However, on closer investigation, these word order alternations reveal some syntactic and semantic differences, hence are not equivalent. I illustrate this point from three perspectives: compatibility with a locative object, position of aspect marker *le* and telicity reading.

2.2. The Syntactic and Semantic differences between the Word Orders

The first difference is the compatibility with a locative NP object. Specifically, in the case of the type 1 DSVC, when the object denotes a location, only the $V_m O V_d$ order is available while the $V_m V_d O$ order becomes ungrammatical.

- (7) a. shang shan lai
ascend mountain come
b. *shang lai shan
ascend come mountain
'come up the mountain (towards the speaker)'

- (8) a. chu xuexiao qu
exit school go
b. *chu qu xuexiao
exit go school
'go out of the school (away from the speaker)'

In the case of the type 3 DSVC, only the $V_m V_d O V_d$ order is grammatical.

- (9) Ta zou shang shan lai.
he walk ascend mountain come
'He walked up the mountain (towards the speaker).'

- (10) *Ta zou shang-lai shan.
he walk ascend-come mountain

- (11) *Ta zou shan shang-lai.
he walk mountain ascend-come

The second difference lies in the distribution of the aspect marker *le*. *le* is traditionally viewed as a perfective aspect marker which occurs after verbs to indicate that the action has been realized.³ *le* is not always allowed to occur after all the verbs in the five orders .

- (12) Ta song (*le) lai le tang. (V_m *le V_d le O)
he send *ASP come ASP soup
'He sent the soup here.'

- (13) Ta song (le) tang lai (le). (V_m le O V_d le)
he send ASP soup come ASP
'He sent the soup here.'

- (14) Ta song (*le) jin (*le) lai (le) tang. (V_m *le V_d *le V_d le O)
he send *ASP enter *ASP come ASP soup

- (15) Ta song (le) tang jin (*le) lai (le). (V_m le O V_d *le V_d le)
he send ASP soup enter *ASP come ASP

- (16) Ta song (*le) jin (le) tang lai (le). (V_m *le V_d le O V_d le)
he send *ASP enter ASP soup come ASP

The third difference is with regard to telicity. Telicity is one of the properties which are used to distinguish Vendler's (1967) four situation types. Basically, a telic event has a

³ The *le* mentioned here only includes aspect marker *le*, namely, the ones immediately following a verb, which I gloss as ASP. Sentence final particle *le* is not relevant to the discussion.

natural finishing point, like ‘eat an apple’, ‘run a mile’. Any event which does not have a natural finishing point is atelic, like ‘walk’, ‘push a cart’. Different word orders appear to result in different telicity. Consider the following:

- (17)a. Ta song lai yi-wan tang, #keshi hai mei dao.
 he send come one-CL soup but still not arrive
 ‘He brought the soup, #but it has not arrived.’
- b. Ta song yi-wan tang lai, keshi hai mei dao.
 he send one-CL soup come, but still not arrive
 ‘He has come with the soup, but it has not arrived.’
- (18)a. Ta song jin-lai yi-wan tang, #keshi zai waimian bei qiang le.
 he send enter-come one-CL soup, but at outside PASS rob ASP
 ‘He brought in a bowl of soup, #but it was robbed outside.’
- b. Ta song yi-wan tang jin-lai, keshi zai waimian bei qiang le.
 he send one-CL soup enter-come but at outside PASS rob ASP
 ‘He is coming sending a bowl of soup, but it was robbed outside.’
- c. Ta song jin yi-wan tang lai, #keshi zai waimian bei qiang le.
 he send enter one-CL soup come but at outside PASS rob ASP
 ‘He brought in a bowl of soup, #but it was robbed outside.’

Take (17) as an example: I add the same assertion ‘but it has not arrived’, which provides a finishing point at the end of the sentences. The $V_m V_d O$ order clashes with this assertion (as in (17a)) while the $V_m O V_d$ order does not (as in (17b)). This indicates that the $V_m V_d O$ order probably entails a finishing point, hence telic while the $V_m O V_d$ order does not necessarily entail such an endpoint, hence atelic. In other words, (17a) indicates that the soup has arrived while (17b) only indicates that he has set off to send the soup, which may have arrived or be on the way. I also use different translations to show this semantic difference. Similarly, the other three word orders also demonstrate different compatibility with such an assertion. The $V_m V_d-V_d O$ order (18a) and the $V_m V_d O V_d$ order (18c) seem to be telic while the $V_m O V_d-V_d$ order (18b) is atelic. This intuition can be further supported by some widely used telicity tests. I will adopt two tests (Dowty 1979) in this paper, which are in-PP test and ‘almost’ test.

In-PP test: basically, in-PP is grammatical with telic events but not compatible with atelic events. For example:

(19) John painted a picture in an hour.

(20) *John walked in an hour.

Applying the in-PP test to the five word orders, we get:

(21) Ta yi-xiaoshi-nei song lai yi-wan tang.
 he one-hour-in send come one-CL soup
 ‘He brought one bowl of soup in one hour.’

- (22) *Ta yi-xiaoshi-nei song yi-wan tang lai.
 he one-hour-in send one-CL soup come
 ‘He came with one bowl of soup in one hour.’
- (23) Ta yi-xiaoshi-nei song jin-lai yi-wan tang.
 he one-hour-in send enter-come one-CL soup
 ‘He brought in one bowl of soup in one hour.’
- (24) *Ta yi-xiaoshi-nei song yi-wan tang jin-lai.
 he one-hour-in send one-CL soup enter-come
 ‘He came in with one bowl of soup in one hour.’
- (25) Ta yi-xiaoshi-nei song jin yi-wan tang lai.
 he one-hour-in send enter one-CL soup come
 ‘He brought in one bowl of soup within one hour.’

The $V_m V_d O$, $V_m V_d-V_d O$ and $V_m V_d O V_d$ orders survive the in-PP test while the $V_m O V_d$ and $V_m O V_d-V_d$ orders do not. Note that (22) and (24) can be grammatical when they mean ‘He will come with the soup in one hour.’, i.e. when ‘in an hour’ means ‘one hour from now’. But the in-PP test cares about the past reading of the in-PP. For this intended reading, (22) and (24) are ungrammatical.

‘Almost’ test: the adverb ‘almost’ has different effects on telic/atelic events.

- (26) John almost painted a picture.

A telic event like (26) gives two readings with ‘almost’: A. John had the intention of painting a picture, but then he changed his mind and did nothing at all. B. John did begin work on the painting and he almost but not quite finished it.

- (27) John almost walked.

An atelic event with ‘almost’ like (27) only has one reading: John did not walk at all.

We use *chadiar* ‘nearly’ as the Chinese counterpart of ‘almost’. I insert it in the five orders and I also add two assertions at the end of each sentence, the former tests the possibility of reading A while the latter tests reading B:

- (28) Ta zuotian chadiar song lai yi-wan tang. Keshi shuiguotou
 he yesterday nearly send come one-CL soup but oversleep
 le jiu mei song. / Keshi song dao banlu bei qiang le.
 ASP then not song. but send arrive halfway PASS rob ASP
 ‘He almost brought one bowl of soup yesterday, but he did not send it because he overslept. / but it was robbed halfway.’
- (29) Ta zuotian chadiar song yi-wan tang lai. Keshi shuiguotou
 he yesterday nearly send one-CL soup come but oversleep
 le jiu mei song. /? Keshi song dao banlu bei qiang le.
 ASP then not send. but send arrive halfway PASS rob ASP
 ‘He almost came with one bowl of soup yesterday, but he did not send it because he overslept. / but it was robbed halfway.’

- (30) Ta zuotian chadiar song jin-lai yi-wan tang. Keshi shuiguotou
 he yesterday nearly send enter-come one-CL soup but oversleep
 le jiu mei song. / Keshi song dao banlu bei qiang le.
 ASP then not song. but send arrive halfway PASS rob ASP
 ‘He almost brought in one bowl of soup yesterday, but he did not send it because he
 overslept. / but it was robbed halfway.’
- (31) Ta zuotian chadiar song yi-wan tang jin-lai. Keshi shuiguotou
 he yesterday nearly send one-CL soup enter-come but oversleep
 le jiu mei song. /? Keshi song dao banlu bei qiang le.
 ASP then not song. but send arrive halfway PASS rob ASP
 ‘He almost came in with one bowl of soup yesterday, but he did not send it because he
 overslept. / but it was robbed halfway.’
- (32) Ta zuotian chadiar song jin yi-wan tang come. Keshi shuiguotou
 he yesterday nearly send enter one-CL soup come but oversleep
 le jiu mei song. / Keshi song dao banlu bei qiang le.
 ASP then not song. but send arrive halfway PASS rob ASP
 ‘He almost brought in one bowl of soup yesterday, but he did not send it because he
 overslept. / but it was robbed halfway.’

As is illustrated above, (28), (30) and (32) allow two readings. In contrast, (29) and (31) definitely allow the first reading while the second reading is odd, if it is not completely impossible.

To sum up, all of the telicity tests show that the $V_m V_d O$, $V_m V_d-V_d O$ and $V_m V_d O V_d$ orders clearly behave like telic events while the $V_m O V_d$ and $V_m O V_d-V_d$ orders are possibly atelic.

2.3. Interim Summary

I summarize the findings we have got for the syntactic and semantic differences between the word orders below:

Table 1: The word orders and their syntactic and semantic properties

Word orders	Compatibility with locative object	Position of <i>le</i>	Telicity
a. $V_m V_d O$	N	$V_m V_d le O$	telic
b. $V_m O V_d$	Y	$V_m le O V_d le$	atelic
c. $V_m V_d-V_d O$	N	$V_m V_d-V_d le O$	telic
d. $V_m O V_d-V_d$	N	$V_m le O V_d-V_d le$	atelic
e. $V_m V_d O V_d$	Y	$V_m V_d le O V_d le$	telic

Based on this table, two questions immediately arise: how can we derive word order (a) to (e)? How can we account for the syntactic and semantic differences observed?

3. Deriving order (a) and (b)

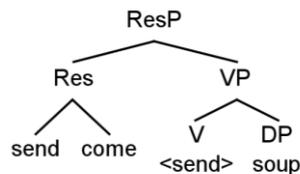
3.1. Proposal

I start from the type 1 DSVC, which involves orders (a) and (b). On the derivation of these two word orders, one solution is to suggest that one order is canonical while the other is

derived from it. The other solution is to say underlyingly, they involve different merge structures. Zou (1994a) and Paul (2008) follow the former approach. However, they did not realise the syntactic and semantic divergence between the two orders, hence their analysis did not and cannot explain the puzzle we have seen. In this paper I try the latter path, which hypothesises two merge structures underlying order (a) and (b).

Specifically, I suggest that order (a) is a resultative construction in which the V_d functions as a resultative predicate. Derivationally, V_m , as the main verb, merges with the object first, projecting VP. Then the V_d is merged as a functional head Res and V will raise to Res, represented in (33).

(33)



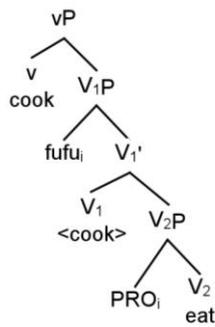
This rough structure can immediately capture the intuition that the perfective marker *le* cannot be inserted in between ‘send’ and ‘come’ because Perf is generally assumed to merge above vP, which is after ‘send’ and ‘come’ have already been adjoined. Besides, (33) also predicts the telic reading of order (a) because if it is a resultative construction, then it is naturally telic. This solution is also not without empirical support. Order (a) behaves very similarly to a Chinese resultative structure, traditionally known as ‘resultative verb compound (RVC)’ (Li & Thompson 1981)⁴. An RVC includes a series of two verbs in which the first one denotes an action and the second one the result from that action. The object is a shared argument of the two verbs and it has to follow the whole verb cluster, as in the case of *da si lang* ‘hit die wolf’ (hit the wolf and the wolf died). RVCs also disallow *le* to intervene between the two verbal elements and are widely known as telic structures. Therefore, it is reasonable to hypothesize that order (a) shares the same structure as an RVC, namely, a resultative structure.

As for order (b) (V_m O V_d), I adopt the approach proposed by Collins (1997) and Paul (2008), which is building up a two VP structure. Collins (1997) investigated the serial verb constructions in Ewe as well as related languages and proposed a structure for SVC, which, against Baker’s three-branch approach to object sharing (Baker 1989), assumes an empty category mediating the relationship between the two verbs in SVC. Briefly, a SVC sentence in Ewe such as (34) is represented by a structure in (35):

- (34) Wo da fufu du.
 3PL cook fufu eat
 ‘They cooked fufu and ate it.’

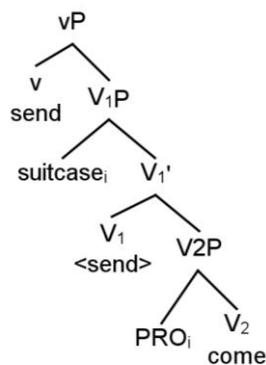
⁴ There is a heated debate on whether RVCs are compounds or phrases and how they are derived in the literature. For a lexical approach, see Yafei Li (1990); for a syntactic approach, see Sybesma (1999). In this paper I adopt a syntactic approach which treats the verbal elements in a resultative structure as two syntactic items rather than two morphemes in a compound.

(35)



Paul (2008) applied Collin's approach to Chinese DSVCs and suggested that the $V_m O V_d$ order is canonical and represents a structure in (36) while the $V_m V_d O$ order is derived from it by raising the lower verb to adjoin to the higher verb.

(36)



I follow Paul in that the $V_m O V_d$ order has the merge structure as in (36). However, I do not suggest that the $V_m V_d O$ order is derived from it, but rather that it is derived as a resultative construction as I mentioned earlier.

3.2. Refining the structure

Previously, I proposed two merged structures (33) and (36) to represent order (a) and (b) respectively. However, this is still not enough to account for the syntactic and semantic characteristics we have seen in Table 1. There are two more issues to address: the syntax of *le* and the syntax of situation aspect.

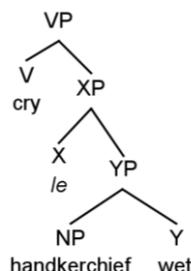
3.2.1. More about the syntax of *le*

In the previous analysis I briefly mentioned that *le* is traditionally called a perfective aspect marker. If this is the case, we would expect it to merge as the Perf head above vP. However, a technical problem would rise in order (b). Recall that order (b) allows two *les* to appear following both verbs at the same time. We cannot expect that to happen if we assume *le* is projected above vP. Hence, we need to reconsider the merge position of *le*. Looking back to the literature about *le*, there are proposals which do not treat *le* as a Perf head above vP. Sybesma (1997), after reviewing various literature on the description of the meaning of *le*, proposed that in syntax, *le* is merged lower than V^0 and serves as a resultative predicate. Therefore, a sentence with both *le* and other resultative predicates indicates an underlying structure: $[_{VP} V [_{XP} X(le) [_{YP} NP Y]]]$. According to him, both XP and YP are small clauses. The verb is complemented by a small clause XP, which is headed by *le*. The head X is complemented by another small clause YP, which is formed by the object and the other

resultative predicate. The semantic relation between XP and YP is that: *le* expresses the state denoted by YP has realized. For instance, (37) has the underlying structure in (38).

- (37) Ta ku shi le shoujuan.
 he cry wet ASP handkerchief
 ‘He cried and the handkerchief got wet.’

(38)



However, to derive the linear order of the three heads, which is ‘cry wet *le*’, Sybesma has to specially assume that *le* is a suffix in the lexicon, hence needs to come last, which is quite stipulative. In this paper, I will adopt his general idea that in a sentence with both *le* and a resultative predicate, *le* signals realization of the result state denoted by the resultative predicate. But I will make some changes to the syntactic position that *le* merges.

3.2.2. Inner aspect

The changes are motivated by the fact that *le* seems to be more closely related to situation aspect rather than viewpoint aspect: it telicizes a resultative phrase which itself telicizes a matrix verb. So it would be nice if we could unify the analysis of *le* with situation types, for which we need Travis' proposal on the syntax of situation aspect.

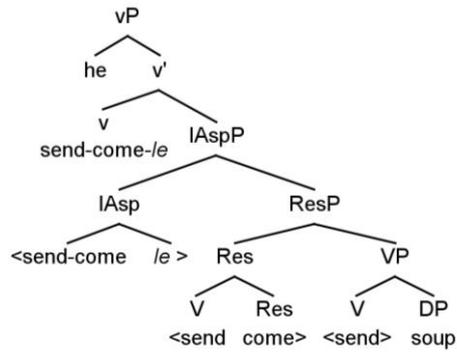
It is well known that the aspect system of human languages consists of two types which are viewpoint aspect and situation aspect (Smith 1997). Travis (2010) proposes that both viewpoint aspect and situation aspect are realized as a head. The situation aspect is realized as a head IAsp (Inner Aspect) which is located between VP and vP while the viewpoint aspect is realized as a head OAsp (Outer Aspect) projected higher above vP. She also suggests that IAsp carries a feature [+/- telic] whose value is computed by the elements within its domain. This value decides the telicity reading.

I adopt Travis' idea that there is an IAsp related to situation aspect between VP and vP. But I propose that what Travis calls the IAsp projection can itself be further articulated, with each smaller projection corresponding to some aspect of telicity. I propose that there are at least two such projections: Res, which hosts the resultative predicate, and IAsp, which can host the aspectual *le* or a null head. And I assume that the IAspP is the place to check telicity features, thus it will always be projected and projected above all the VPs if there is more than one VP in the clause, whereas ResP is only projected when a resultative predicate occurs. What I proposed above is represented in (39).

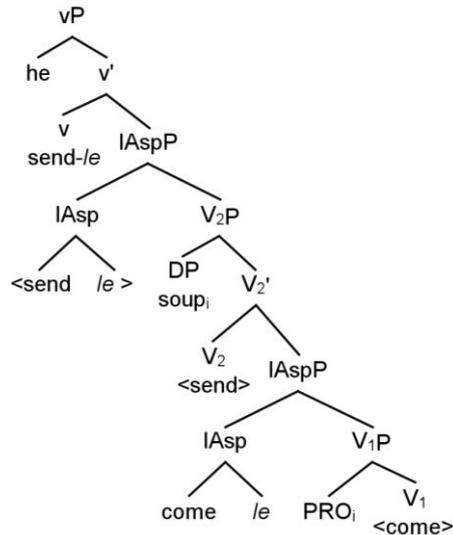
- (39) [_{IAspP} IAsp(*le*/ ϕ) [_{ResP} Res [_{VP} V DP]]]

V will raise to Res and then the cluster V-Res further raises to IAsp via head movement. In minimalist terms, Res has a strong [uV*] feature, which triggers V to raise and adjoin to it. And similarly, IAsp also has a [uV*] feature which forces V along with Res to adjoin to it. In this way we can derive the right surface order without assuming *le* is a suffix. (40) represents the derivation of order (a) and (41) represents the derivation of order (b).

(40) Derivation of order (a)



(41) Derivation of order (b)



Based on this we can also derive the telicity reading. Recall that we assume the IAsp head could be null or *le*. When *le* is not presented, order (a) yields telic reading while order (b) atelic reading. Hence, we can assume that when the head of IAsp is null, the null IAsp head bears a feature [Tel:]. When V along with Res moves to the IAsp, the feature [Res] on Res will value the null head as [Tel:Res], so the sentence yields telic reading. When the ResP is not projected, only V raises to IAsp, the [Tel:] feature will remain unvalued and by default it will be interpreted as atelic. In contrast, when the head of IAsp is *le*, the overt IAsp head *le* bears a feature [Tel:Res]. No matter if there is a Res moving to IAsp, the value on IAsp will not be changed because the feature already has a [Res] value. This is supported by the fact that when *le* occurs, the sentence is always telic no matter whether there is a resultative predicate.

4. Deriving word order (c)-(e)

Now we move on to the word orders (c)-(e) which are repeated below:

c. V_m V_d-V_d O (*song jin-lai tang* ‘send enter-come soup’)

d. V_m O V_d-V_d (*song tang jin-lai* ‘send soup enter-come’)

e. V_m V_d O V_d (*song jin tang lai* ‘send enter soup come’)

Here we encounter a dilemma on whether to treat the V_d cluster in orders (c) and (d) as a compound or two independent verbs. In the descriptive literature, the V_d cluster such as *jin-lai* is viewed as a compound without question. Besides, Table 1 clearly shows that order (c)

behaves like order (a) and, order (d) like order (b) in terms of the position of *le* and the telicity reading. If the V_d cluster is really a compound, then order (c) and (d) would be equal to (a) and (b) respectively, which naturally predicts their similarity in *le*'s distribution and telicity. However, it would be hard to explain why order (e) exists if we assume the V_d cluster is a compound.

Further evidence shows that the split V_{ds} in order (e) clearly show phrasal properties while the V_d cluster in orders (c)-(d) behaves like a single word: when the double V_d cluster is not separated by an NP, *le* can never appear in the middle. But when the cluster is separated by an NP, *le* can be inserted between two V_{ds} :

- (42)a. Ta song jin (*le) lai shu.
 he send enter (ASP) come book
 b. Ta song jin le shu lai.
 he send enter ASP book come

To solve this dilemma, I suggest that the V_d cluster in orders (c)-(d) is indeed a compound while in order (e) there is no compound at all. Instead, the two V_{ds} are two lexical items from the lexicon, namely two heads in the syntax.

It may be more economical to argue for a coherent analysis which treats all the V_{ds} in orders (c)-(e) as independent verb heads. However, this first counters native speakers' intuition that the V_d cluster, such as *jin-lai*, is a single word rather than two words. More importantly, empirical data shows that the V_d cluster in orders (c)-(d) is indeed a compound verb. Compare the following pairs:

- (43)a. *Ta song le.
 he send ASP
 '*He sent.'
 b. *Ta jin le.
 he enter ASP
 '*He entered.'

- (44)a. *Ta song lai le.
 he send come ASP
 '*He brought.'
 b. Ta jin-lai le.
 he enter-come ASP
 'He came in.'

Both *song* and *jin* are transitive when they are used alone, so if there is no overt object following them, the sentences become ungrammatical, as in (43).⁵ However, when they are combined with the directional verb *lai*, the verb string *song lai* disallows no object (as in (44a)) while *jin lai* is grammatical with no object (as in (44b)). If our analysis in Section 3 is correct, the string *song lai* involves two heads in which the V head *song* needs a complement, so if there is no object occurring, the uninterpretable [uD] feature on V will not be checked. If *jin-lai* also involves two heads, it should also require an object, but this is not the case. This suggests that *jin-lai* enters syntax as a whole word, more specifically, an intransitive verb. Therefore, I suggest that all the V_d clusters in orders (c)-(d) should be compounds, which

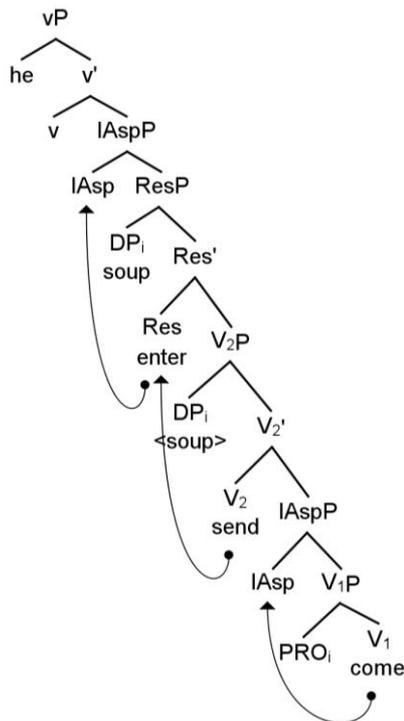
⁵ Sentences in (43) are definitely ungrammatical on the non-pro drop reading. They are acceptable when there is an implicit object, but if there is an implied object, we should expect a PRO in syntax which still means they need an internal argument.

means orders (c)-(d) would bear the same structure and derivation process as orders (a)-(b) respectively.

On the other hand, in order (e), we assume the two V_d s separated by an NP are different lexical items, hence the lexical integrity condition is not violated. Then how are they represented in syntax? The occurrence of *le* gives us some clue. Remember that in order (e), *le* can appear after the two V_d s but not after the V_m . Take (42b) for example, *le* can occur after *jin* and after *lai*, but not after *song*. Following my analysis in Section 3, *le* will be attached to either a V head or a Res head. Then we have reason to believe that here *jin* and *lai* are either a V head or Res head.

For *lai*, it is fairly clear that it can only be a V head because Res must be merged above a VP and attached to V after V moves to it, but there is no other V heads sticking to *lai* on surface, therefore *lai* is a V head. As for *jin*, if it is a V head, we should expect that *le* can follow both *song* and *jin* because according to my hypothesis, there is an IAspP above each VP. However, as I just mentioned *le* cannot occur after *song*, hence, *jin* cannot be a V head. However, if *jin* is a Res head, the position of *le* is well predicted because the main verb *song* will move to Res (*jin*) first and then they together move to IAsp head, which leads to the surface order V-Res-*le*. Therefore, I suggest that in order (e), the first V_d is a Res head while the second V_d is a V head. Then the merge structure underlying order (e) turns out to be a double VP structure in which the higher VP, projected by V_m , has a ResP above it projected by a non-deictic V_d whereas the lower VP is projected by a deictic V_d . (45) is the tree diagram for the merge structure of word order (e) with the arrows showing the head movement.

(45) Derivation of order (e)



If both of the IAsp heads are filled by *le*, we will derive the right linear order: *song jin le tang lai le* ‘send enter *le* soup come *le*’. And the telic reading of order (e) is also bore out because the higher IAsp is valued as [Tel:res] by the Res, thus yielding telic interpretation.

5. More remarks on the Locative object

The theory we have developed so far explains the position of *le* in word orders (a)-(e) as well as the telicity value of each order. The remaining puzzle is why some of the orders can take a locative object while some cannot.

We first look at orders (a) and (b). Basically, order (a) does not allow a locative object while (b) does. The examples are repeated below:

- (46) a. *ta shang lai shan.
 he ascend come mountain
 b. Ta shang shan lai.
 he ascend mountain come
 ‘He came up the mountain.’

Following the theory we developed in Section 3, (46b) can be derived without question. Specifically, it involves two VPs which are ‘ascend mountain’ and ‘lai’, as in (47). After the head movement (V to IAsp), we get (46b).

- (47) $[_{VP} \text{he}_i [_{V'} V [_{IAspP} \text{IAsp} [_{V2} \text{mountain} [_{V2'} \text{ascend} [_{IAspP} \text{IAsp} [_{V1P} \text{PRO}_i \text{come}]]]]]]]]]$

Nonetheless, the theory we have so far cannot predict the ungrammaticality of (46a). Nothing in (48) prevents the derivation to converge successfully.

- (48) $[_{VP} \text{he} [_{V'} V [_{IAspP} \text{IAsp} [_{ResP} \text{come} [_{VP} \text{ascend mountain}]]]]]]]$

To solve this, I assume that Res head bears a [uD] feature which triggers the object NP to move to [Spec, Res]. So the derivation for *song lai tang* ‘send come soup’ would be modified as (49):

- (49) $[_{ResP} \text{soup} [_{Res'} \text{come} [_{VP} \text{send} \langle \text{soup} \rangle]]]]]$

Semantically, (49) would be interpreted as ‘...sent the soup and the result is the soup came’. However, if the object is a locative object, although syntactically we can still move the LocO to [Spec, Res], as in (50) which will result in (46a) after head movement, we would get an anomalous interpretation: He ascended the mountain, as a result, the mountain came.

- (50) $[_{VP} \text{he} [_{V'} V [_{IAspP} \text{IAsp} [_{ResP} \text{mountain} [_{Res'} \text{come} [_{VP} \text{ascend} \langle \text{mountain} \rangle]]]]]]]]]$

In other words, the unacceptability of (46a) is because order (a) combined with a locative object yields infelicitous reading. To derive the intended reading, we have to resort to order (b).

Now we move to orders (c)-(e). Orders (c)-(d) disallow locative object while order (e) is grammatical with locative object. I repeat the examples below:

- (51) a. *Ta zou jin-lai fangjian. (order (c))
 he walk enter-come room
 b. *Ta zou fangjian jin-lai. (order (d))
 he walk room enter-come

- c. Ta zou jin fangjian lai. (order (e))
 he walk enter room come
 'He walked into the room.'

Recall we argued before that orders (c)-(d) are equal to orders (a)-(b) respectively. Then the ungrammaticality of (51a) is expected. However, (51b) is not expected because order (b) should be grammatical with a locative object. Of course we can argue that in (51b), the V_m *zou* is intransitive, hence it should not take an object. Nevertheless, even if we add a preposition after *zou* or swap *zou* with a transitive verb, it is still ungrammatical:

- (52) *Ta zou dao fangjian jin-lai.
 He walk to room enter-come

- (53) *Ta jin fangjin jin-lai.
 he enter room enter-come

I suggest that this is because the semantics of the complex direction compounds such as *jin-lai* indicates a location which is known to both the speaker and listener or can be inferred from the context. To see this more clearly, consider the following scenarios:

Scenario 1: Xiaoming walked up the mountain where Xiaohong and her friends are standing on. Then Xiaohong can say either of the following sentences to her friends:

- (54)a. Ta zou shang shan lai le.
 he walk ascend mountain come ASP
 b. Ta zou shang-lai le.
 he walk ascend-come ASP
 'He walked up (the mountain).'

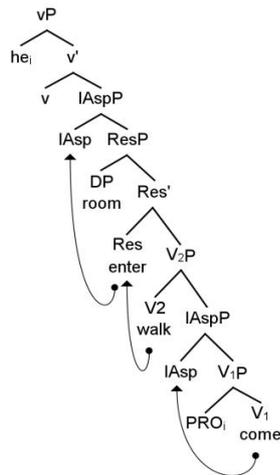
Scenario 2: Xiaoming swam across the river. Xiaohong and her friends are standing on the bank watching. Then Xiaohong can say either of the following sentences to her friends:

- (55)a. Ta you guo he qu le.
 he swim cross river go ASP
 b. Ta you guo-qu le.
 he swim cross-go ASP
 'He swam across (the river).'

In other words, the conceptual location semantically incorporated in the direction compounds such as *shang-lai* and *guo-qu* cannot co-occur with an overt locative NP, hence, when a directional compound is used, no more overt locative object is allowed.

Finally, order (e) with a locative object should be derived as in (56), where the LocO is merged directly in the [Spec, Res]:

(56)



6. Conclusion

In this paper, I have investigated five word orders shown by the directional serial verb construction in Mandarin Chinese. I show that these word orders reveal many syntactic and semantic differences, including their compatibility with a locative object, the position of *le* and their telicity interpretation. To derive the five word orders and account for the puzzles we identified, I propose that order (a) represents a resultative structure while order (b) a double VP structure. I further argue that the V_d cluster in orders (c) and (d) is a compound, so (c) and (d) should share the same derivation as (a) and (b) respectively. Finally, order (e) involves a double VP structure in which the higher VP has a ResP projected above it.

References

- Baker, M. C. (1989). Object sharing and projection in serial verb constructions. *Linguistic inquiry*, 513-553.
- Collins, C. (1997). Argument sharing in serial verb constructions. *Linguistic inquiry*, 461-497.
- Dowty, D. R. (1979). Word meaning and Montague grammar: The semantics of verbs and times in generative semantics and in Montague's PTQ. Springer Science & Business Media.
- Li, C. & Thompson, S. (1981). *Mandarin Chinese: A functional reference to Chinese grammar*. Berkeley: University of California Press.
- Li, Y. (1990). On VV compounds in Chinese. *Natural language & linguistic theory* 8(2), 177-207.
- Paul, W. (2008). The serial verb construction in Chinese: A tenacious myth and a Gordian knot. *The Linguistic Review* 25, 367-411.
- Smith, C. S. (1997). *The parameter of aspect*. Springer Science & Business Media.
- Sybesma, R. (1997). Why chinese verb-le is a resultative predicate. *Journal of East Asian Linguistics* 6(3), 215-261.
- Sybesma, R. P. (1999). *The mandarin VP*. Kluwer Academic Publishers.
- Travis, L. (2010). *Inner aspect: The articulation of VP*. Springer Science & Business Media.
- Vendler, Z. (1967). *Linguistics in philosophy*. Cornell University Press Ithaca.
- Zou, K. (1994a). Directional verb compounds in Chinese. In Proceedings of the 30th Annual meeting of the Chicago Linguistic Society 1, 443-457.

Zhishuang Chen
Department of Language and Linguistic Science, University of York
1Apollo Street
YO10 5AP
UK

zc643@york.ac.uk

TWO TYPES OF STATES IN KOREAN: EXPERIMENTAL LI EVIDENCE*

JIYOUNG CHOI
(University of Nantes)

Abstract

This paper establishes that there are two classes of states in Korean: pure states vs. so-called inchoative states. We argue that pure states (e.g. *celmta* ‘young’, *pikonhata* ‘tired’) describe a homogeneous state without an inherent transition, while inchoative states (e.g. *nulhta* ‘old’, *cichita* ‘tired’) describe a state with an inherent transition into that state. Since these two classes of states have different inherent aspectual properties, when combined with the perfect marker *-ess*, they yield different temporal readings. That is, with pure states, *-ess* yields an anterior reading, while with inchoative states, it yields a simultaneous reading with respect to the utterance time. This paper further discusses whether Korean children aged from four to six can assign different temporal readings of *-ess* to pure states and inchoative states.

1. Introduction

In the literature, the most well-known aspectual classification is proposed by Vendler (1967) where four aspectual classes are distinguished: states, activities, accomplishments and achievements (see also Dowty 1979, Smith 1997, Carlson 1994 many others). In the present study, we are particularly interested in the class of states. States (e.g. *know*, *believe*, *love*, *be happy*) describe certain properties or continuous situations which hold all throughout a given period of time. They do not entail any change or transition.

States in Korean have been argued to divide into two sub-classes: pure states vs. so-called inchoative states¹ (cf. Chung 2005, Choi 2015), as given in (1).

- (1) a. PS: *celmta* ‘young’, *pikonhata* ‘tired’, *pisushata* ‘similar’, *nalssinhata* ‘thin’, ...etc
b. INS: *nulhta* ‘old’, *cichita* ‘tired’, *talmta* ‘alike’, *maluta* ‘thin’, ...etc.

At first glance, it seems that both classes of predicates appear to describe certain properties or states of individuals or objects. For instance, both the pure state predicate *celmta* ‘young’ and the inchoative state predicate *nulhta* ‘old’ describe a property of age.

When combined with a past marker *-essess* (cf. Chung 2005), both pure states and inchoative states yield a typical stative reading, as illustrated in (2a-b).

- (2) a. Sue-ka caknyeny/hantongan-unnalssinha-essess-ta.²
Sue-NOM last.year/for.a.while-TOP thin-PAST-DEC
‘Sue was thin last year/for a while.’ [pure state]

* I would like to thank the audience at the Second Asian and European Linguistic Conference (AE-Link2) and, in particular, Hamida Demirdache and Orin Percus for helpful comments and discussion. Any remaining errors or infelicities are entirely my own.

¹ Note that inchoative states have been observed in several other languages such as *Skw̃w̃w̃* (Bar-el 2005), *Səncáθən* and Japanese (Kiyota 2008) as well as Spanish reflexive psychological verbs (Marín & McNally 2011).

² The following abbreviations are used in this paper: ACC=accusative, DEC=declarative, INCHO=inchoative, NOM=nominative, PAST=past, PFCT=perfect, PRES=present, TOP=topic.

- b. Sue-ka caknyeney/hantongan-unmalu-essess-ta.
 Sue-NOM last.year/for.a.while-TOP thin-PAST-DEC
 ‘Sue was thin last year/for a while.’ [inchoative state]

In (2a-b), the pure state predicate *nalssinha* ‘thin’ and the inchoative state predicate *malu* ‘thin’ respectively describe that the state of Sue’s being thin happened with a temporal duration prior to utterance time and no longer holds at utterance time. As such, they give rise to a past stative reading.

However, when combined with a perfect marker *-ess* (cf. Chung 2005, Choi 2015), pure states and inchoative states crucially yield different temporal readings. With pure states, *-ess* yields a past reading, allowing only modification by past time adverbials such as *cinancuey* ‘last week’, as shown in (3a). On the contrary, with inchoative states, *-ess* yields an on-going result state reading, allowing only modification by present time adverbials such as *cikum* ‘now’, as illustrated in (3b).

- (3) a. Sue-ka **cinancuey**/***cikum** apfu-ess-ta.
 Sue-NOM last.week/now sick-PFCT-DEC
 ‘Sue was sick last week.’ / *‘Sue is sick now.’ [pure state]
- b. Sue-ka **cikum**/***cinancuey** hwana-ss-ta.
 Sue-NOM now/last.week angry-PFCT-DEC
 ‘Sue is angry now.’ / *‘Sue was angry last week.’ [inchoative state]

The fact that pure states and inchoative states do not pattern together with respect to the perfect marker *-ess* suggests that these two classes of states do not have the same temporal structures.

This paper aims to establish that inchoative states do not belong to the class of pure states (i.e. typical stative predicates), but rather they constitute a distinct class of predicates that the standard classification does not include. We provide novel experimental evidence from Korean child language for the claim.

This paper is structured as follows. In Section 2, we invoke a set of diagnostics allowing us to distinguish pure states and inchoative states. Based on the results of these diagnostics, we also provide a proposal on the lexical meaning of the two classes of states in Korean. In Section 3, we discuss temporal interpretation of the two classes of states. Reviewing the fact that the perfect marker *-ess* attached to pure states and inchoative states yields different temporal readings, we demonstrate that an anterior (i.e. past) and a simultaneous (i.e. on-going present) readings of the two classes of states are not expressed in the same way. Section 4 deals with the issue of whether Korean children can distinguish the two classes of states in temporal contexts. Specifically, we present an experiment designed to investigate whether children can assign different temporal readings of the perfect marker *-ess* to the two classes of states. Finally, Section 5 summarizes the main points of this paper.

2. Pure states vs. Inchoative states

2.1. Diagnostics distinguishing the two classes of states

Pure states and inchoative states can be distinguished with respect to several diagnostics. In this section, we particularly invoke three diagnostics (cf. Choi 2015 for other diagnostics): (i) the present marker *-nun/-Ø*; (ii) the inchoative marker *-e ci*; (iii) the addition of a punctual adverbial clause.

2.1.1. Present marker *-nun/-Ø*

First, pure states and inchoative states can be morphologically distinguished with respect to the present marker *-nun/-Ø*. It has been traditionally argued that verbal predicates in Korean can be distinguished from adjectival (or non-verbal) predicates in that they take different present morphemes. Verbal predicates take an overt present marker *-nun* (or its allomorph *-n*), while adjectival predicates take a zero morpheme *-Ø* (cf. Soh 1995, Han 1996, Chung 1999 many others). This is illustrated in (4).

- (4) a. *Minho-nun cikum sakwa-lul mek-nun/*Ø-ta.*
 Minho-TOP now apple-ACC eat-PRES-DEC
 ‘Minho is eating/eats an apple now.’ [verbal predicate]
- b. *Minho-nun cikum haksayng-i-Ø/*nun-ta.*
 Minho-TOP now student-be-PRES-DEC
 ‘Minho is a student now.’ [nominal predicate]

As can be seen in (4), the verbal predicate *mek* ‘eat’ takes the overt present marker *-nun*, while the nominal predicate *haksayng* ‘student’ does not take it.

Let us now turn to the two classes of states. Consider the examples given in (5).

- (5) a. *Minho-ka pikonha-Ø/*nun-ta.*
 Minho-NOM tired-PRES-DEC
 ‘Minho is tired.’ [pure state]
- b. *Minho-ka cichi-n/*Ø-ta.*
 Minho-NOM tired-PRES-DEC
 ‘Minho is getting tired.’ [inchoative state]

As can be seen in (5a), pure states do not take the overt present marker *-nun*. Specifically, in (5a), the bare form of the pure state predicate *pikonha* ‘tired’ describes a state of Minho’s being tired holding at the utterance time. On the contrary, inchoative states felicitously combine with the overt present marker *-nun* like other verbal predicates, as shown in (5b). The contrast in (5a-b) suggests that inchoative states should be distinguished from pure states since pure states are adjectival predicates, while inchoative states are verbal predicates.

2.1.2. Inchoative marker *-e ci*

Second, the two classes of states show different behavior with respect to the overt inchoative marker *-e ci*. In Korean, the inchoative verb entailing a change-of-state is derived by the addition of the inchoative morpheme *-e ci* (roughly translated as ‘BECOME state’) (cf. Chung 2005, Joo 2008, Lim 2010). To illustrate, consider the examples given in (6).

- (6) a. *Ku maktayki-ka kil-ess-ta.*
 that stick-NOM long-PFCT-DEC
 ‘That stick was long.’
- b. *Ku maktayki-ka kil-e ci-ess-ta.*
 that stick-NOM long-INCHO-PFCT-DEC
 ‘That stick became long(er).’

(Lim 2010)

The pure state *kil* ‘long’ in (6a) describes a state of being long without entailing a change-of-state. In (6b), the inchoative marker *-e ci* marking the addition of a BECOME operator to the event structure, affixes to the pure state predicate *kil* ‘long’ and as such, it gives rise to an inchoative interpretation where a transition from NOT BEING LONG to BEING LONG takes place and the described state starts to hold at the utterance time (i.e. That stick is long at the utterance time).

However, the inchoative marker *-e ci* cannot take inchoative states as its argument, as shown in (7).

- (7) a. Sue-ka hwana-ess-ta.
 Sue-NOM angry-PFCT-DEC
 ‘Sue got angry.’
 b. Sue-ka hwana-***e ci**-ess-ta.
 Sue-NOM angry-INCHO-PFCT-DEC

In (7a), the inchoative state *hwana* ‘angry’ on its own gives rise to an inchoative reading where the transition from NOT BEING ANGRY to BEING ANGRY occurs and the described (resultant) state holds at the utterance time (i.e. Sue is angry at the utterance time). In (7b), the inchoative state cannot felicitously co-occur with *-e ci*, unlike pure states. The ungrammaticality of adding *-e ci* to the inchoative state in (7b) suggests that inchoative states are inherently inchoative, that is, they have an initial zero-marked BECOME operator in their lexical meaning, blocking the addition of another, overt BECOME operator.

2.1.3. Addition of a punctual adverbial clause

Third, pure states and inchoative states do not pattern together with respect to the addition of a punctual adverbial clause. The addition of a punctual adverbial clause can induce three readings according to the aspectual properties of predicates (cf. Bar-el 2005), as given in (8).

- (8) a. INCEPTIVE (inchoative) reading: the described eventuality in the main clause begins at the same time as the event described by the punctual adverbial clause
 b. MEDIAL (overlapping) reading: the described eventuality takes place simultaneously with the event described by the punctual adverbial clause
 c. CULMINATING reading: the described eventuality ends simultaneously with the event described by the punctual adverbial clause

If the inceptive (inchoative) reading is the only available reading, then the matrix predicate contains an initial boundary (i.e. the onset of the described eventuality) in its meaning. Let us first consider what happens with inchoative states.

- (9) Juno-nun [ku sosik-ul tul-ess-ul ttay] hwana-ss-ta.
 Juno-TOP that news-ACC hear-PFCT-when angry-PFCT-DEC
 ‘Juno was angry when he heard that news.’ [inchoative state]
 a. ✓Juno was not angry before, but he became angry because of the news.’
 b. ✗Juno was already angry when he heard the news.’

The sentence (9) contains an inchoative state predicate *hwana* ‘angry’ in the main clause co-occurring with a punctual clause. The utterance (9) can be accepted only under an inceptive context (9a): at the time where the punctual event occurs, the state described by the predicate in the main clause begins simultaneously. A medial reading (9b) is not available for the sentence. It suggests that inchoative states in Korean refer to the onset of the state they are associated with, as part of their meaning.

However, pure states do not show the same behaviour as inchoative states with respect to this diagnostic. Consider the following example.

- (10) Juno-nun [nay-ka cenhwahay-ss-ul ttay] aphu-ess-ta.
 Juno-TOP I-NOM call-PFCT-when sick-PFCT-DEC
 ‘Juno was sick when I called him.’ [pure state]
 a. ✓Juno was already sick when I called him.
 b. ✗Juno was not sick before, but he got sick when I called him.’

The sentence (10) containing a pure state predicate *aphu* ‘sick’ can be accepted only under a medial context (10a). It does not describe a change-of-state from NOT BEING SICK to BEING SICK as in (10b). The unavailability of an inceptive reading for the sentence (10) suggests that, unlike inchoative states, pure states in Korean describe a state without referring to the onset (i.e. an initial transition) of the state they are associated with.

2.2. Proposal on the lexical meaning of the two classes of states

Based on the results of the diagnostics discussed so far, we claim that pure states describe a homogeneous state without involving an inherent transition or a change, while inchoative states describe a state with an inherent transition (i.e. BECOME) into that state. Specifically, the transition entailed in inchoative states is represented as the ONSET of the state that they are associated with (following Bar-el 2005). Consequently, pure states and inchoative states do not have the same event representation, as provided in (11).

- (11) a. (pure) States: $\lambda e.P(e)$ (cf. Dowty 1979, Rothstein 2004)
 b. Inchoative states: $\lambda e.\exists e_1\exists e_2[e=S(e_1\sqcup e_2) \wedge (\text{BECOME}(P))(e_1) \wedge P(e_2)]$ (cf. Bar-el 2005)

In (11a), pure states have a simplex event structure in that they contain only a state (e), which is durative. On the other hand, in (11b), inchoative states have a complex event structure in that they contain an initial BECOME sub-event (e_1) followed by a (resultant) state (e_2) which is durative. As such, pure states yield a state meaning, while inchoative states yield an inchoative meaning.

With the proposal in mind, let us now review the temporal readings of the perfect marker *-ess* that we briefly observed in Section 1.

3.2. Temporal readings of the perfect marker *-ess*

As we saw in (3a-b) discussed in Section 1, pure states and inchoative states crucially yield the different temporal readings. Assuming that the suffix *-ess* is a perfect marker (cf. Lee 1991, Han 1996, Chung 2005) yielding an anterior (i.e. past) reading or a simultaneous (i.e. on-going result state) reading with respect to utterance time (i.e. reference time in simple clauses), we argue that the distribution of the temporal readings of *-ess* is determined by the event structure of predicates with which it occurs (cf. Choi 2010).

With pure states (PS), *-ess* yields an anterior (ANT) reading, allowing only modification by past time adverbials as shown in (3a) which is repeated in (12a) below. Note that, as illustrated in (5a) and (12b) above, a simultaneous reading (SIM) of pure states is expressed by the null present marker $-\emptyset$.

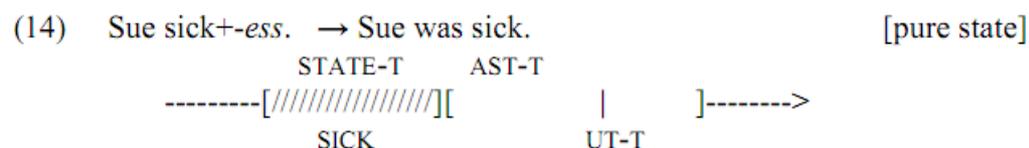
- (12)a. Sue-ka **cinancuey/*cikum** aphu-ess/essess³-ta.
 Sue-NOM last.week/now sick-PFCT/PAST-DEC
 ‘Sue was sick last week.’ / *‘Sue is sick now.’ [PS+*-ess/essess*: ANT reading]
- b. Sue-ka **cikum/*cinancuey** aphu- \emptyset -ta.
 Sue-NOM now/last.week sick-PRES-DEC
 ‘Sue is sick now.’ / *‘Sue was sick last week.’ [PS+ $-\emptyset$: SIM reading]

With inchoative states (INS), *-ess* yields a simultaneous reading, allowing only modification by present time adverbials as shown in (3b) which is repeated in (13a) below. Note that, as illustrated in (13b), an anterior reading of inchoative states is expressed by the real past marker *-essess* (cf. Chung 2005).

- (13)a. Sue-ka **cikum/*cinancuey** hwana-ss-ta.
 Sue-NOM now/last.week angry-PFCT-DEC
 ‘Sue is angry now.’ / *‘Sue was angry last week.’ [INS+*-ess*: SIM reading]
- b. Sue-ka **cinancuey/*cikum** hwana-ssess-ta.
 Sue-NOM last.week/now angry-PAST-DEC
 ‘Sue was angry last week.’ / *‘Sue is angry now.’ [INS+*-essess*: ANT reading]

To account for the different temporal readings of the perfect marker *-ess*, I adopt Demirdache & Uribe-Etxebarria (2007)’s temporal syntax where Tense/T^o and Aspect/Asp^o are spatiotemporal ordering predicates projecting their time denoting arguments in the syntax: (i) T^o orders its external argument (utterance time (UT-T) in matrix clause) relative to its internal argument (assertion time (AST-T); Klein 1994); (ii) Asp^o orders its external argument (AST-T) relative to its internal argument (event time⁴ (EV-T)). D & U-E argue that the perfect is a spatio-temporal predicate with the meaning of after. More specifically, *present* orders the UT-T **within** the AST-T and then, *perfect* orders the AST-T **after** the EV-T.

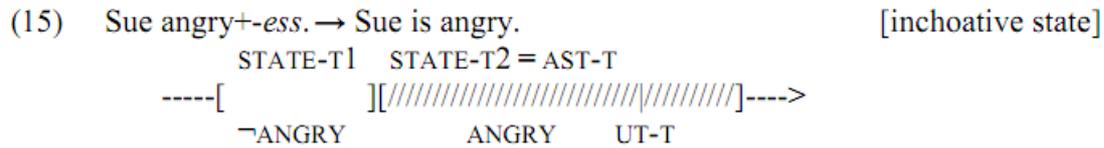
Adopting D & U-E’s account, we propose that the perfect marker *-ess* orders **immediately after** the time interval in its immediate scope, i.e. the EV-T (or STATE-T). First, as we claimed, pure states have a simplex event structure (cf. (11a)). When attached to a pure state predicate, *-ess* orders the AST-T immediately after the STATE-T and as such, an anterior reading is generated (i.e. the STATE-T is construed as past-shifted relative to the UT-T). This is illustrated in (14).



³ Notice that the double form *-essess* affixed to pure states patterns with the perfect marker *-ess* in yielding an anterior interpretation.

⁴ The internal argument of Asp^o (i.e. VP) can be assigned either a simplex event structure (EV-T), or a complex event structure consisting of two sub-events (e.g. process and result state for telic predicates) according to the predicate type under consideration. Each sub-event projects the time argument defining its running time (e.g. process-time (EV-T1) and result state-time (EV-T2)). See D & U-E (2007) for the details.

Inchoative states have a complex event structure of consisting of two sub-events (cf. (11b)). When attached to an inchoative state predicate, *-ess* orders the AST-T immediately after the interval defining the STATE-T1 of the inchoative state, thus focusing the time defining the result state (STATE-T2) and as such, a simultaneous reading is generated. This is illustrated in (15).



Thus, the simultaneous reading of inchoative states suffixed by *-ess* is further evidence for the presence of a complex event structure of inchoative states (compared to that of pure states) in Korean.

Now, we ask ourselves an acquisition question of whether Korean children are aware of the fact that the combination between the perfect marker *-ess* and the two classes of states yields different temporal readings. We explore this question by an experiment to which the next section is devoted.

4. Experiment

4.1. Participants

The study involved thirty Korean children in total (ten 4-year-olds (from 4;3 to 4;8 with a mean of 4;7), ten 5-year-olds (from 5;0 to 5;7 with a mean of 5;3) and ten 6-year-olds (from 6;0 to 6;11 with a mean of 6;4)) and twenty Korean adults as the control group. All children and adults who participated in the task were native Korean speakers.

4.2. Design

A truth-value judgment task (TVJT; cf. Crain & Thornton 1998 many others) was carried out to investigate whether Korean children can distinguish the two classes of states in the anterior and the simultaneous temporal contexts. Four experimental conditions were constructed in a 2x2 design with predicate type (PS vs. INS) and context type (ANT vs. SIM) as factors. In the ANT context where the target state occurred prior to utterance time and is no longer true at utterance time, both PS and INS were proposed. Likewise, in the SIM context where the target state holds at utterance time, the same two types of states were proposed.

Each participant was presented with sixteen target items (four items per condition) interspersed with sixteen distractors and control items, for a total of thirty-two items. Children watched stories acted out by means of an animated PowerPoint slide show on a laptop screen. At the end of each story, a puppet made a statement in answer to the lead-in question asked by the experimenter. The participants' task was to judge whether the puppet's statement was true or false in the given context. They were also asked to volunteer a target sentence to describe the given context. This follow-up production task was done to ascertain whether they accepted or rejected the test sentences for the expected reasons. (16)-(19) give examples of the experimental stimuli translated into English.

(16) Condition 1: PS+*-ess* in the ANT context

Scenario: Sue caught a cold. Sue is very sick with fever and her mother worries about Sue. So, she takes Sue to hospital to see the doctor. In the hospital, Sue got an injection and took medicine. The next morning, Sue got over her cold and she is feeling well.

Lead-in question: How was Sue?

Test sentence: Sue-ka aphu-ess-eyo.
Sue-NOM sick-PFCT-DEC
'Sue was sick.'

Expected answer: Yes

(17) **Condition 2: INS+-ess in the ANT context**

Scenario: Ppororo is very angry. What could we do to calm down Ppororo? Look! Eddy who is Ppororo's best friend came to play with Ppororo. Eddy says: "Hey, Ppororo! Why are you so angry? Calm down and let's play a game with me." Now, Ppororo is fine. And this is thanks to Eddy!

Lead-in question: How was Ppororo?

Test sentence: Ppororo-ka hwana-ss-eyo.
Ppororo-NOM angry-PFCT-DEC
'Ppororo is angry.'

Expected answer: No

(18) **Condition 3: PS+-ess in the SIM context**

Scenario: Juno likes to play with his friends, painting with his hands. Juno has paint on his hands and makes handprint on a paper. Juno wants to show the handprint paper to his mom. But, look! His hands are too dirty. Juno goes to the bathroom and washes his hands with soap and water. Now, his hands are clean.

Lead-in question: How are Juno's hands?

Test sentence: Juno-uy son-i kkaykkuha-ess-eyo.
Juno-GEN hand-NOM clean-PFCT-DEC
'Juno's hands were clean.'

Expected answer: No

(19) **Condition 4: INS+-ess in the SIM context**

Scenario: The weather is very nice today. Piglet who likes sunlight goes out and takes a walk. Suddenly, the sky is filled with rain clouds. It is raining! Piglet didn't bring his umbrella. So, he gets all wet in the rain.

Lead-in question: How is Piglet?

Test sentence: Akitoeci-ka phi-ey cec-ess-eyo.
Piglet-NOM rain-in wet-PFCT-DEC
'Piglet is wet in the rain.'

Expected answer: Yes

4.3. Results

4.3.1. Overall results

Table 1 below presents the overall results of the control group and children.

Table 1. TVJT: overall results of adults vs. children (% of acceptance)

	ANT context		SIM context	
	PS	INS	PS	INS
Adults (n=20)	100	15	1.25	97.5
Children (n=30)	99.17	67.5	49.17	100

First of all, Korean adults, the control group, correctly assigned the ANT reading to PS and the SIM reading to INS: they correctly accepted PS (100% of acceptance), but rejected INS (15% of acceptance) in the ANT context. Likewise, they correctly accepted INS (97.5% of acceptance), but rejected PS (1.25% of acceptance) in the SIM context. Paired-samples *t*-tests revealed that the control group significantly distinguished INS from PS in both the ANT context ($t(19) = 12.350, p < .001$) and the SIM context ($t(19) = -35.184, p < .001$).

Overall, children accepted PS in the ANT context (99.17% of acceptance) and INS in the SIM context (100% of acceptance). However, their rejection rate of PS in the SIM context (50.83%) and that of INS in the ANT context (32.50%) do not seem to be enough to conclude that they assigned different temporal readings of *-ess* to each of the two classes of states.

In what follows, we break down the results by age groups to examine whether there is variation in adult-like and non-adult-like behavior across age groups.

4.3.2. Results by age groups

Table 2 below presents the results of children by age groups.

Table 2. TVJT: results of children by age groups (% of acceptance)

	ANT context		SIM context	
	PS	INS	PS	INS
4-year-olds (n=10)	100	97.5	77.5	100
5-year-olds (n=10)	97.5	72.5	57.5	100
6-year-olds (n=10)	100	32.5	12.5	100

First, as shown in Table 2 above, 4-year-olds did not show the expected performance in that they accepted both PS and INS in both the ANT context (100% of acceptance for PS, 97.5% of acceptance for INS) and the SIM context (77.5% of acceptance for PS, 100% of acceptance for INS). These high acceptance rates suggest that 4-year-olds were not sensitive to different temporal readings of *-ess* on PS and INS. In other words, they did not make a significant distinction between PS and INS in both the ANT context ($t(9) = 1.000, p = .343$) and the SIM context ($t(9) = -2.077, p = .068$).

5-year-olds performed better than 4-year-olds, but still did not show the expected target performance. Like the control group, 5-year-olds accepted PS in the ANT context (97.5% of acceptance) and INS in the SIM context (100% of acceptance). However, unlike the control group, they also accepted PS in the SIM context (57.5% of acceptance) and INS in the ANT context (72.5% of acceptance). Note that their unexpected acceptance rate of PS in the SIM (57.5%) is lower than that of 4-year-olds (77.5%). This relatively lower acceptance indicates that, unlike 4-year-olds, 5-year-olds distinguish PS from INS in the SIM context ($t(9) = -3.157, p = .012$). Likewise, 5-year-olds' unexpected acceptance rate of INS in the ANT context (72.5%) is lower than that of 4-year-olds (97.5%), but this acceptance rate per se is still high. In other words, 5-year-olds tended to accept both PS and INS in the ANT context, suggesting that they did not significantly distinguish the two classes of states in this context ($t(9) = 2.236, p = .052$). Thus, 5-year-olds' behavior can be summarized as follows: they distinguish the two classes of states at least in the SIM context, but not in the ANT context.

Let us now turn to 6-year-olds, who patterned with the adults control group. They correctly accepted PS in the ANT context (100% of acceptance) and INS in the SIM context (100% of acceptance). On the other hand, they mostly rejected INS in the ANT context (32.5% of acceptance) and PS in the SIM context (12.5% of acceptance), as expected. As such, 6-

year-olds significantly distinguished PS from INS in both the ANT context ($t(9) = 4.521, p = .001$) and the SIM context ($t(9) = -8.720, p < .001$), like adults.

The question arises as to why 5-year-olds distinguished the two classes of states in the simultaneous context, but not in the anterior context?

4.3.3. Children's patterns

Following their *yes/no* answers to the puppet's statements, children were asked to describe *what happened in the story?*. In order to achieve a better understanding of Korean children's knowledge of the distinction between the two classes of states, we considered children's production data in relation to the comprehension results. According to their behavior, we divided the children into three groups. The first group includes children who correctly assigned the relevant temporal reading of *-ess* to the two classes of states, as shown in (20).

(20) **Target-like pattern (overall: 30% of children)**

Context:	ANT		SIM	
	PS	INS	PS	INS
Comprehension:	yes	no	no	yes
Production:	-ess	-essess	-∅	-ess

These children have perfectly acquired temporal interpretation of both pure states and inchoative states. That is, they have plausibly acquired the target-like semantics of the relevant tense/aspect markers (i.e. *-ess* vs. *-essess* vs. *-∅*). As a result, they distinguish the two types of states in both the anterior context and the simultaneous context respectively.

The second group includes children who totally failed to assign the relevant temporal reading of *-ess* to the two classes of states, as illustrated in (21).

(21) **Non-target-like pattern (overall: 16.67% of children)**

Context:	ANT		SIM	
	PS	INS	PS	INS
Comprehension:	yes	yes	yes	yes
Production:	-ess	-ess	-ess	-ess

These children have not acquired temporal interpretation of both pure states and inchoative states, yet. They are unable to distinguish the two types of states in both the anterior context and the simultaneous context. Crucially, their knowledge of *-ess* is not adult-like because it seems to be specified as having either an anterior or a simultaneous reading regardless of the type of predicate.

The third group includes children who were accurate with pure states, but were inaccurate with inchoative states, as shown in (22). We refer the children of this group as 'partially target-like children'.

(22) **Partially target-like pattern (overall: 53.33% of children)**

Context:	ANT		SIM	
	PS	INS	PS	INS
Comprehension:	yes	yes	no	yes
Production:	- <i>ess</i>	- <i>ess</i>	- \emptyset	- <i>ess</i>

These children have fully acquired temporal interpretation of pure states. That is, they have acquired the target-like semantics of *-ess* and $-\emptyset$ (i.e. *-ess* is specified as yielding an anterior reading, and $-\emptyset$ is specified as yielding a simultaneous reading of pure states). However, they seem to have some problem with inchoative states because they accepted inchoative states combined with *-ess* in both the anterior and the simultaneous contexts. This leads them to make a distinction between pure states and inchoative states at least in the simultaneous context, but not in the anterior context.

Overall, 30% of the child participants showed target-like pattern, 16.67% of children non-target-like pattern, and 53.33% of children partially target-like pattern.

4.4. Discussion

In the previous section, the overall results of the TVJT showed that Korean children have some difficulties with the temporal readings of the perfect marker *-ess* on pure states and inchoative states. This seemed to be the case for 4-year-olds who seem to make no significant distinction between the two classes of states in both the anterior and the simultaneous context. It also seemed to be the case for 5-year-olds who made a significant difference between the two classes of states in the simultaneous context, but not in the anterior context. 6-year-olds, however, behaved like adults. Then, the question arose as to why some children distinguished the two classes of states in the simultaneous context, but not in the anterior context? In other words, why did children incorrectly accept inchoative states combined with *-ess* in the anterior context?

On the basis of the children's patterns that we found in light of the follow-up production data and the breakdown comprehension results, we argue that the children who showed the partially target-like pattern (cf. (22)) did not acquire the real past marker *-essess* which is the only possible form for an anterior interpretation of inchoative states. Due to the lack of *-essess*, these children are likely to generalize inchoative states combined with *-ess* for both an anterior and a simultaneous temporal interpretations.

Another question that arises as to why some younger children (i.e. non-target-like children) incorrectly accepted both pure states and inchoative states in both the anterior context and the simultaneous context. It could be either because they have a problem with the meaning of the perfect marker *-ess*, or because they have a problem with the distinction between pure states and inchoative states. To seek a plausible answer to this question, however, further experimental investigation is needed (*in progress*).

5. Conclusion

In this paper, we have distinguished the two classes of states in Korean, i.e. pure states and inchoative states. We argued that a pure state describes a property without referring to an inherent transition or change (i.e. typical stative predicates), while an inchoative state describes a property as well as the inherent transition (i.e. BECOME sub-event) into the described property. Since these two classes of states do not have the same properties, when

they are combined with the perfect marker *-ess*, they do not show the same behavior. With pure states that have a simplex event structure, *-ess* yields an anterior reading, while with inchoative states that have a complex event structure, *-ess* yields a simultaneous reading. Then, we provided novel experimental evidence from Korean child language for the distinction between pure states and inchoative states in temporal contexts. Our L1 acquisition results revealed that, overall, the distinction between the two classes of states seems to be acquired by children at the age of five and six. In particular, we found that 5-year-olds distinguished the two classes of states at least in the simultaneous context. We suggest that the reason why they did not make such distinction in the anterior context is related to the lack of the past marker *-essess* in their grammar. Thus, the results illustrate that pure states are acquired earlier than inchoative states in child language. This is not so surprising since pure states correspond to typical stative predicates that most languages have, while inchoative states correspond to a particular class of predicates that some languages have.

References

- Bar-el, L. (2005). *Aspectual Distinctions in Skwxwúmesh*. Ph.D. dissertation, University of British Columbia.
- Carlson, R. (1994). *A grammar of Supyire*. Berlin: Mouton de Gruyter.
- Choi, J.Y. (2010). Inchoative states and the perfect marker *-ess* in Korean, paper presented: 5th Workshop of the European Research Net in Linguistics, University of the Basque Country (UPV-EHU).
- Choi, J.Y. (2015). On the universality of aspectual classes: inchoative states in Korean. In: E. Labeau & Q. Zhang (eds.), *Taming the TAME Systems, Cahier Chronos 27*, 123-135.
- Chung, K.S. (1999). *Deictic and anaphoric tense in Korean: A two-tiered approach*. MA thesis, Simon Fraser University.
- Chung, K.S. (2005). *Space in Tense: The interpretation of Tense, Aspect, Evidentiality and Speech Act in Korean*. Ph.D. dissertation, Simon Fraser University.
- Crain, S. & Thornton, R. (1998). *Investigations in Universal Grammar: A Guide to Experiments on the Acquisition of Syntax and Semantics*. Cambridge, MA: MIT Press.
- Demirdache, H. & Uribe-Etxebarria, M. (2007). The syntax of time arguments. *Lingua: International Review of General Linguistics*, Vol. 117.2, 330-366. Oxford, United Kingdom: Elsevier.
- Dowty, D. (1979). *Word meaning and Montague Grammar*. Dordrecht Reidel.
- Han D.-W. (1996). *Kwukeyu sicey yenkwu* (A study on Korean tense). Seoul: Kwukehakhoy Taehaksa.
- Joo, H.R. (2008). *Agentivity of passives and inchoatives in second language learners of English and Korean*. Ph.D. dissertation, University of Hawaii.
- Kiyota, M. (2008). *Situation aspect and viewpoint aspect from Salish to Japanese*. Ph.D dissertation, University of British Columbia.
- Klein, W. (1994). *Time in language*. London: Routledge.
- Kee, H.-S. (1991). *Tense, aspect and modality: A discourse-pragmatic analysis of verbal affixes in Korean from a typological perspective*. Ph.D. dissertation, University of California.
- Lim, D.S. (2010). The syntax of inchoatives: *-e ci*, event structure and scalarity. In: J. Yeon & J. Kiaer (eds.), *Selected papers from the 2nd European conference on Korean linguistics*. Munchen: Lincom Europa.
- Marín, R. & McNally, L. (2011). Inchoativity, change of state, and telicity: Evidence from Spanish reflexive psychological verbs. *Natural Language and Linguistics Theory* 29, 467-502.

- Rothstein, S. (2004). *Structuring Events: A study in the semantics of aspect*. Wiley-Blackwell.
- Smith, C. (1997). *The Parameter of Aspect*. Dordrecht: Kluwer Academic Press.
- Soh, S.-O. (1995). *Tense and aspect in Korean*. Honolulu, Hawaii: University of Hawaii Press.
- Vendler, Z. (1967). *Linguistics in philosophy*. Cornell University Press.

Jiyoung Choi
Linguistics Laboratory of Nantes (LLING)
Chemin de la Censive du Tertre
44312 Nantes
France

choi.jiyoung1003@gmail.com

*LICENSING DEFINITENESS: A SPANNING ACCOUNT OF NOUN PHRASE INTERPRETATION IN MANDARIN CHINESE AND CANTONESE**

DAVID HALL

(Queen Mary University of London)

Abstract

In this paper I provide a novel account of a number of facts related to restrictions on the interpretation of Mandarin Chinese and Cantonese noun phrases that have been discussed extensively in the literature. The approach that I take eschews a more standard head-movement approach, and implements a ‘spanning’ account, in which multiple heads in an extended projection can be spelled out as a single morpheme. I show that a spanning account of the interpretation of different noun phrase configurations in conjunction with a condition on the licensing of a definite D head in the extended nominal projection captures all of the facts, and is also compatible with data from a classifier language, Nuosu Yi, which exhibits definite bare nouns and also a definite determiner.

1. Introduction and proposal

I present here an analysis of constraints on noun phrase interpretation in Mandarin Chinese (MC) and Cantonese making use of an approach to the syntax-morphology interface which relies on the notion of morphological ‘spans’: sequences of heads in an extended projection that can be the target for insertion of a single morpheme (Abels & Muriungi 2007, Svenonius 2012, Adger 2013, a.o.). I show that such an approach leads to a parsimonious account of the facts, and also gives us an explanation for an interesting quirk of the languages, whereby adjectives in a high structural position in the DP force a definite interpretation. I also show that the spanning story allows us to avoid a word order problem which is faced by a standard head-movement account.

The presentation is structured as follows: in section 2 I present noun phrase interpretation facts in MC, focusing on the special case of modifier position in subsection 2.1. In section 3 I propose an account of those facts based on the idea of morphological spans, and expand my discussion to Cantonese, showing how the account can capture the data there too. In section 4 I briefly give an alternative head-movement account, and show that there is one case in particular where that account gets the facts wrong. In section 5 I discuss some cross-linguistic implications of the system that I set up, before concluding in section 6.

2. Noun phrase interpretation in Mandarin Chinese

The availability of a definite or indefinite reading for a noun phrase in MC depends on a number of factors, including syntactic position and also the configuration of elements internal to the noun phrase. Bare nouns in object position can be definite or indefinite, and are

* I thank David Adger, Hagit Borer, Daniel Harbour, Fangfang Niu, Tom Stanton, and the audience of AE-Link 2 for valuable feedback on this paper. I would also like to thank Cherry Lam, Hong Liu, Fanfang Niu, Chen Wang, Joanna Wat, Panpan Yao and Annette Zhao for their judgements.

number neutral. Classifier-Noun (CI-N) sequences are obligatorily singular, and are unambiguously indefinite. Numeral-CI-N sequences are unambiguously indefinite, and are singular or plural depending on the numeral (singular if it is *yi* ‘one’, plural otherwise). See Cheng & Sybesma (1999, 2005) for an in depth discussion of these facts.¹

The preverbal subject position is restricted to definite noun phrases (Huang, Li & Li 2009, a.o.), and so CI-N and Numeral-CI-N sequences are degraded in this position (1). Bare nouns are acceptable, as are noun phrases with a demonstrative, which must have a definite interpretation in this position (2).²

(1) a. *??san-ge xuesheng chi-le dangao*
 three-CL student eat-LE cake
 Intended: ‘Three students ate the cake.’

b. *??ge xuesheng chi-le dangao*
 CL student eat-LE cake
 Intended: ‘A student ate the cake’

(2) a. *gou chi-le dangao*
 dog eat-LE cake
 ‘The dog ate the cake’

b. *nei-zhi gou chi-le dangao*
 that-CL dog eat-LE cake
 ‘That dog ate the cake’

In table 1 I summarize the distributional facts.

Table 1: Noun phrase interpretation in MC

Noun Phrase Config	Definite	Indefinite	Number
N	Yes	Yes	Neutral
CI-N	No	Yes	Singular
Num-CI-N	No	Yes	Singular/plural

2.1. Modifier position and interpretation

Phrasal adjectives canonically appear immediately pre-nominally, accompanied by the general modifier marker *de*.³ Modified bare nouns maintain their ambiguity with respect to

¹ I do not discuss the availability of a generic interpretation for bare nouns here, but recognise that any theory of the interaction between syntax and interpretation in the noun phrase must be able to account for generic readings as well as definite/indefinite contrasts. This I leave for future research.

² I gloss *le* simply as LE, not wanting to commit to any specific analysis of its function, which has been widely debated in the literature. It has been analysed as a perfective aspect marker (Huang et al 2009), and as a resultative predicate (Sybesma 1997).

³ I leave aside here discussion of bare modifiers which do not occur with *de*, as their distribution is far more restricted, and their behaviour is somewhat idiosyncratic and subject to interspeaker variability. See Yang (2005), chapter 6, for in depth discussion of the differences between the two types of modifier. I also put aside discussion of the exact function of *de*, as nothing I say hinges on this. I therefore gloss *de* simply as DE.

definiteness. When a numeral and classifier are present, the unmarked order of elements is Numeral > Cl > Adj > N, and the phrase is unambiguously indefinite. Adjectives can also appear in a marked pre-numeral position:

- (3) a. *san-zhi huangse-de gou* (unmarked pre-nominal position)
 Three-CL yellow-DE dog
- b. *huangse-de san-zhi gou* (marked pre-numeral position)
 yellow-DE three-CL dog
 ‘three yellow dogs’

The characteristics of noun phrases which have a marked pre-numeral adjective have been discussed in quite some detail in the literature. I summarise the properties of these ‘high modifiers nominals’ (HMNs, nominals with the order Adj-Numeral-Cl-N) in (4).⁴

- (4) a. *HMN in postverbal subject position (Zhang 2015);
 b. *HMN in *you* existentials (Huang 1982, Lu 1998);
 c. *HMN after non-referential *ta* (Lin & Zhang 2006);
 d. HMN acceptable in pre-verbal subject position (my informants);
 e. *Extraction of N from HMN (Zhang 2015);
 f. HMN exhibits forced ‘wide-scope’ reading with respect to quantifiers (my informants);
 g. HMN exhibits uniqueness/inclusiveness presupposition (my informants).

Characteristics (a) to (d) show that HMNs pattern with definite and strongly quantified noun phrases. Characteristic (e) shows that HMNs exhibit the kind of resistance to N extraction that definite DPs have been argued to exhibit (Diesing 1992). Property (f) shows that HMNs do not behave like indefinites in MC with respect to their scope relative to other quantificational elements (see Huang 1982), and (g) shows that they exhibit the same kind of uniqueness/inclusiveness presupposition that definite descriptions do (see, e.g., Lyons 1999, Elbourne 2012). Each of these properties is the opposite of what holds for low modifier nominal (LMNs), which are obligatorily interpreted as being indefinite. In table 2 I summarize the interpretational facts related to modified noun phrases:

Table 2: Noun phrase interpretation in MC (with modifiers)

Noun phrase config	Definite	Indefinite
Adj -N	Yes	Yes
Cl-Adj-N	No	Yes
Numeral-Cl-Adj-N (LMN)	No	Yes
Adj-Numeral-Cl-N (HMN)	Yes	No

These facts are intended to highlight a significant contrast in MC noun phrase interpretation: Numeral-Cl-N sequences are obligatorily indefinite, and remain so when a modifier is introduced in the canonical low position, but HMNs are obligatorily definite. In the remainder of this paper, I will argue that these facts, along with a cross-linguistic

⁴ My informants all self-identify as native speakers of Mandarin. They are all in roughly the same age group (mid-twenties).

generalization relating to definiteness and noun phrase configuration, can be understood to follow from the interaction of a constraint on the licensing of a definite D head with the way that the syntax interfaces with morphology through spell-out. I also show that the position of the adjective relative to a definite bare noun cannot be explained through a head movement analysis of the facts, but that an account which utilizes the notion of morphological spans is able to capture the data.

3. A ‘spanning’ account

Jenks (2012) identifies and discusses a crosslinguistic generalization that holds across a number of classifier languages which allow bare nouns in argument position. The generalization takes the form of two one-way entailments: (i) if a classifier language has bare nouns which can be interpreted as definite, then CI-N phrases will be unambiguously indefinite; (ii) if a classifier language has CI-N phrases which can be interpreted as definite, then bare nouns will be unambiguously indefinite. This is summarized in (5):

- (5) a. Bare N [\pm def] \rightarrow CI-N [$-$ def] (Type A language)
 b. CI-N [\pm def] \rightarrow Bare N [$-$ def] (Type B language)

The generalization holds across a number of Sino-Tibetan and Austroasiatic classifier languages including Hmong, Cantonese, Mandarin, Min, Wu⁵, and Vietnamese.

Jenks proposes that this pattern can be understood as resulting from a parametric difference in the way that parts of the functional structure of the DP spell out in the two different language types. The proposal relies on the notion of morphological SPAN, where more than one head in an extended projection can be spelled out as a single morphological unit (see Abels & Muriungi 2003, Svenonius 2012, a.o.). To summarize, Jenks proposes that Type A languages (such as Mandarin) have a span over [$D_{[def]}$ [CI [N]]] which spells out as a noun (giving a definite interpretation for bare nouns), and Type B languages (such as Cantonese) have a span over [$D_{[def]}$ [CI]] which spells out as a classifier (giving a definite interpretation for CI-N sequences). Type A languages do not have the span available in Type B languages, and vice versa. This means that there is no way for a definite D head to spell out as part of a classifier in a language like Mandarin, and there is no way for a definite D head to spell out as part of a bare noun in a language like Cantonese.

In the next subsection I give a brief explanation of how spanning works, and then show how we can expand on Jenks’s insight to account for the entire range of interpretations available in Mandarin and Cantonese.

3.1. Spelling out Spans

I adopt the assumptions in (6) about the syntax-morphology interface, which constitute a spell-out system which follows Williams (2003) and Adger (2013), and which is similar in principle to Svenonius (2012):^{6,7}

⁵ CI-N can give rise to a definite interpretation in Wu only where there is a change in tone on the classifier. See Cheng & Sybesma (2005) for discussion.

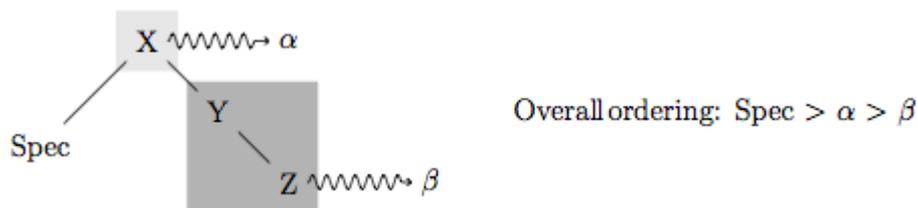
⁶ See also Dékány (2011) for a related system.

⁷ The assumption (6d) is a potential locus of parametric variation; thus it would be more accurate to state that this is universal spell-out system with an additional language specific parameter setting.

- (6) a. A single morpheme can spell out more than one head in a (continuous) extended projection line of heads; i.e., spell-out targets spans.
 b. All heads must spell out as part of a span (possibly as a trivial, one head span), or have an overt specifier.
 c. Specifiers spell out to the left of heads.
 d. Spans are ordered by their relative height ('higher than' = 'left of').

In what follows, all trees are given as 'telescoped' representations, in the sense of Brody (2000), as such representations make complement lines clearly visible. Categories in the projection line stand for both a head and its phrasal projection, with specifiers being leftward branches coming off of those heads. An illustration of the system of spell-out that I adopt is given in (7). Spans are shown through shading, and wiggly lines represent the spell-out of those spans:

(7)



Here a span from Z to Y is spelled out as a single morpheme β (6a). This morpheme is spelled out low, in the position of Z. The (trivial) span over X is spelled out as the morpheme α . α and β are ordered by assumption (6d), which says that the higher span spells out to the left of the lower span, giving $\alpha > \beta$. 'Spec' is spelled out to the left of the span over X (6c), giving Spec > α . Thus the total ordering is Spec > α > β . I am using Spec as a place holder for some other potentially complex unit; whatever its internal composition, it will appear preceding α in the example in (7). If the element labelled 'Spec' was instead a specifier of Y, then the ordering would be α > Spec > β .

3.2. The syntax and semantics of the extended nominal projection

I make some specific assumptions about the syntax of nominals here, and about the semantic contribution of the different heads in the extended nominal projection (extended projection in the sense of Grimshaw 2005).⁸ The semantic function of Cl is to impose a join complete semi-lattice structure over the property that the noun denotes, identifying atoms and sets of those atoms which have the property P. The nature of the semi-lattice means that both atoms and pluralities are included: the structure produced through the merger of Cl is interpreted as number neutral. The Num head introduces a variable over that structure, and imposes further restrictions on the lattice through the bivalent feature [\pm atomic].⁹ [+atomic] identifies only the atoms in the semi-lattice, excluding non-atoms, and [-atomic] identifies all of the non-atoms, excluding the atoms. This gives us the singular/plural distinction. For

⁸ See Borer (2005, 2013) and Adger (2013) for more recent thorough explorations of the idea of extended projections.

⁹ The feature [\pm atomic] adopted here is taken from Harbour (2014).

example, the structure [Num_[-atomic] [Cl [N]]] would give us all of the non-atomic sets of things with the property given by N. A numeral is introduced in the specifier of a functional head Q. If there is a mismatch between the feature on Num and the value given by the numeral, then the result is semantically incoherent (although I assume that the syntax is able to build such a structure).

I assume that an overt classifier signals the projection of a Num head in MC, and that the default feature value on Num is [+atomic], as Cl-N sequences are obligatorily singular. In the case of bare nouns, a Num head is projected but with no feature specifying atomicity (or lack thereof); in this case Num's only job is to introduce a variable over the structure, and we get a number neutral interpretation (which includes the whole semi-lattice, namely all of the atoms and sets of atoms). I do include a Cl head in the extended projection of even a bare noun, as it has semantic content (i.e., bare nouns can be count and denote atoms and sets composed of those atoms).

The variable introduced by Num can be existentially closed at the VP level (Heim 1984, Diesing 1992, a.o.), or can be bound by an iota operator in the D head, which is the topic of the next subsection.¹⁰

3.3. Licensing definiteness

I take D to be the locus of definiteness; D introduces the iota operator which binds the variable introduced by Num. Where D merges, the noun phrase is interpreted as definite, but only if the structure meets a licensing condition on definiteness.¹¹ The disjunctive licensing condition on a D head, which I take to hold cross-linguistically, is as follows:¹²

- (8) A D head is licensed iff either
- a. the head is spelled out as part of a span with other heads, or
 - b. the head is spelled out itself (a trivial span), as a definite article, or
 - c. a phonologically overt element merges in the specifier of D.

The intuition behind this condition is that there has to be some kind of morphological 'flagging' of the D position for it to be interpreted. The way that I have worded the proposal above is potentially misleading: it appears that I am positing a morpho-phonological licensing condition on the syntactic structure, which would seem not to be in the spirit of the Y-model (and associated models) of the architecture of the grammar, in which phonological and interpretative processes are post-syntactic, and on separate branches which do not interact.

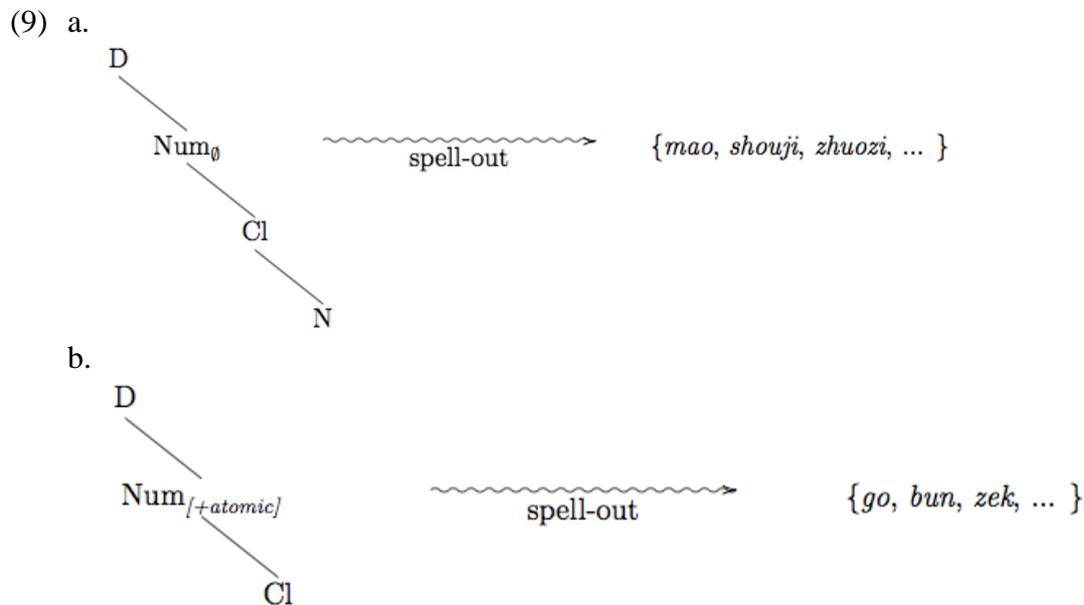
However, the proposal is instead that there are restrictions on what phonological outputs a speaker has available to her; a speaker has a list of learned phonological outputs given a specific morpho-syntactic structure, and if the speaker does not have an independent phonological output for D (i.e., an article), then the only way that a definite interpretation can

¹⁰ I do not discuss generic interpretations here but also assume that the possibility of a Generic operator binding the variable is available.

¹¹ As noted above, the lack of a D projection means that the variable introduced by Num can be bound through existential closure, giving rise to an indefinite interpretation.

¹² Ultimately I hope to derive this constraint from a more general constraint on ALL HEADS, i.e., a constraint which states that any head in the syntax is licensed if it is associated with some phonological material (either a specifier or a span). For the purposes of this paper, however, the more specific constraint given in (8) will have to suffice.

be achieved is through signalling the presence of a definite D in some other way. The syntax can presumably build whatever structure it wants, but if there is no output then the result is unacceptability. Available spell-outs of spans must simply be learned and stored in a list. For example, in MC, we have the information in (9a) stored, and in Cantonese we have the information in (9b) stored (more on the specifics of this in sections 3.4 and 3.5):



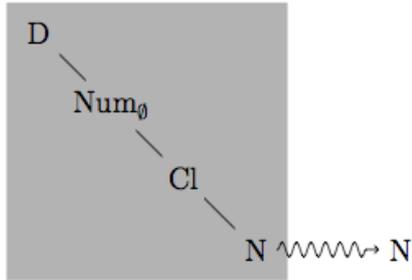
Necessarily then, ‘flagging’ possibilities will be limited by the kind of input that the child gets. The reason that Cl-N configurations in MC can never be definite is assumed to be that there was nothing in the input that would lead a learner to posit the span in (9b).¹³ This way of conceptualizing the constraint requires that we assume that the structure of nominals is universal and fixed across languages: when attempting to navigate the hypothesis space of possible interpretations for nominals, the child has access to a set of constraints which allow it to close in on the appropriate representations that match the phonology (and the meaning) of the input. Thus we ultimately place the locus of cross-linguistic variation in the lexicon, and in the ways in which syntactic structure maps to a morpho-phonological output stored therein.

3.4. Structure of the noun phrase in MC

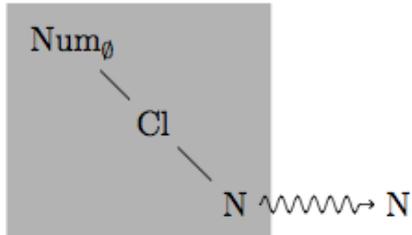
Here I present the structures associated with each of the different noun phrase configurations available in MC, and show how the constraint on D discussed above allows us to explain the constraints on interpretation on each type of noun phrase. Spans across heads are highlighted in grey, wiggly lines show spell-out, and the result of spell-out is given as N or Cl, which are stand-ins for phonological strings which are members of the sets of elements that a speaker normally refers to as a ‘noun’ or ‘classifier’, respectively (e.g., N might be *gou* ‘dog’, or *tuoxie* ‘slippers’, and Cl might be *ge* (general classifier), or *duo* (classifier for flowers, clouds)):

¹³ This would lead us to imagine that a diachronic shift from Cl-N definites to bare N definites, and vice versa, is quite possible, if not likely. I leave this to future research.

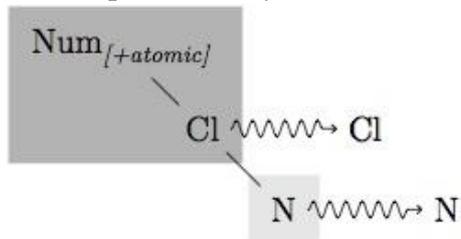
(10) *Bare noun (definite)*



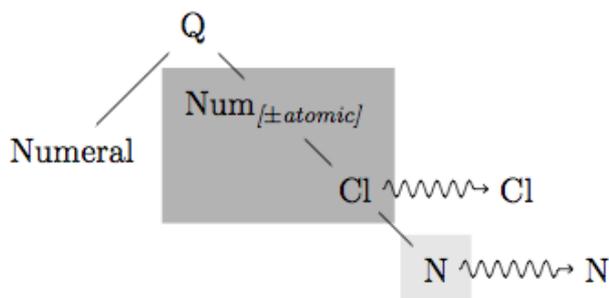
(11) *Bare noun (indefinite)*



(12) *Cl-N sequence (indefinite)*



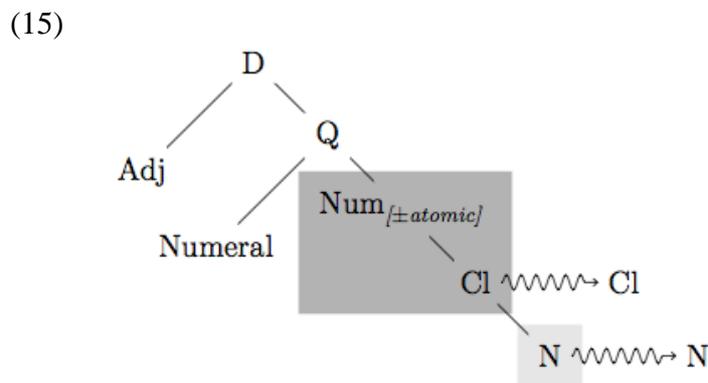
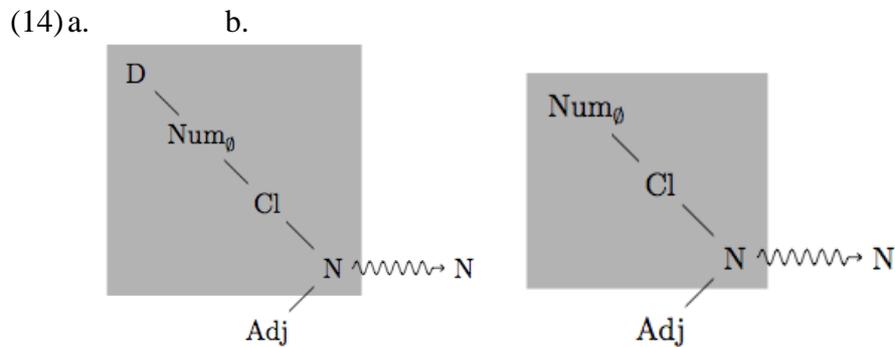
(13) *Numeral-Cl-N sequence (indefinite)*



In the case of Cl-N sequences, a definite interpretation is never available because there is no span from Cl to D through Num_[+atomic] which spells out as Cl in Mandarin. A span from N up to D is only available when Num is empty (Num_∅), forcing a number neutral interpretation for bare nouns. A span across Cl and Num is possible when Num is either [+atomic] or [−atomic], spelling out as a classifier, capturing the fact that the presence of a numeral can give rise to either singular or plural interpretation, depending on the number.

When an adjective is merged in the low position, immediately preceding the noun, the available interpretations are the same as with unmodified bare nouns. This is because the spans spell out low in N, and specifiers spell out to the left of heads. (14a) shows a definite Adj-N sequence, (14b) an indefinite Adj-N sequence. In the case of HMNs, the adjective is

merged in spec D,¹⁴ licensing the definite head, and forcing a definite interpretation (15). If no D projects, there is no high merge site for the adjective. Where D hosts an adjective, it meets the licensing criterion in (8c), and the iota operator binding the individual variable introduced in Num means that a definite interpretation is forced. I assume that the low adjective merge site is either in the spec of a dedicated functional head (Cinque 1994, et seq.),¹⁵ or is in spec N, but chose the latter in the graphical representation of the structure for clarity's sake.



We have seen how each of the available interpretations for different configurations of argument nominal can be derived. I now extend the spanning account to Cantonese, an example of a ‘type B’ language.

3.5. Extending the account to Cantonese

Cantonese has the following interpretational constraints on the noun phrase:

Table 3: Noun phrase interpretation in Cantonese

Noun Phrase Config	Definite	Indefinite	Number
N	No	Yes	Neutral
Cl-N	Yes	Yes	Singular ¹⁶
Num-Cl-N	No	Yes	Singular/plural

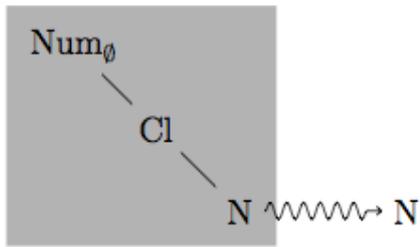
¹⁴ For reasons of space, I leave aside the question of whether this is a base merge position or a position resulting from a movement operation (from the lower, canonical position for adjectives).

¹⁵ Spans from N upward in this case would have to include that functional head.

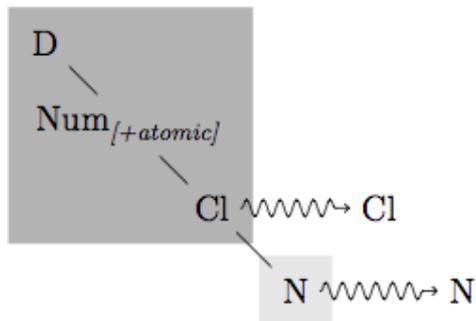
¹⁶ A *di*-N sequence is plural, but unfortunately I cannot discuss *di* here for reasons of space.

We can account for the Cantonese facts in the same way as with Mandarin, assuming the following structures (with spans highlighted in grey).

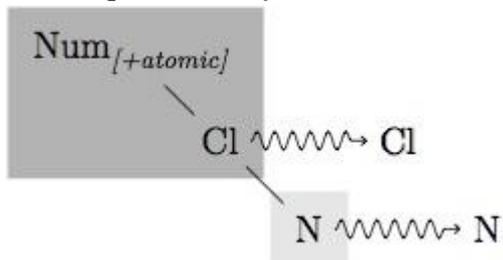
(16) *Bare noun (indefinite)*



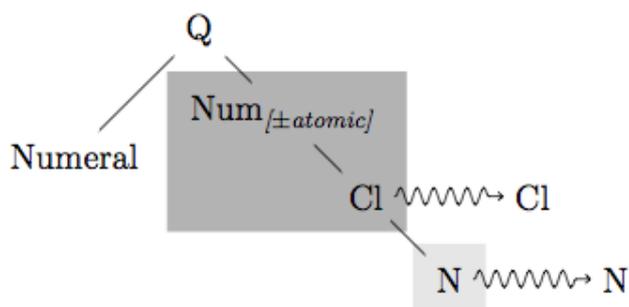
(17) *Cl-N sequence (definite)*



(18) *Cl-N sequence (indefinite)*



(19) *Numeral-Cl-N sequence (indefinite)*



The indefinite Cl-N, indefinite bare noun and Numeral-Cl-N sequences have the same structure as we saw for Mandarin. The definite Cl-N sequence, however, involves a span from Cl to $\text{Num}_{[+atomic]}$ to D, giving rise to a definite reading. This is the span that is not available in MC.

Summarizing everything discussed so far, we have the generalizations given in table 4 about spans across functional heads in the domain of the noun phrase in Mandarin and Cantonese (on the assumption that Q blocks any span past Num):

Table 4. Spanning availability in Mandarin and Cantonese

	Up to Num _o	Up to Num _[+atomic]	Up to D
Span from N	Both	Neither	Mandarin only
Span from Cl	Neither	Both	Cantonese only

We see that the difference between the two languages lies in how far up certain spans are allowed to go. Spans starting at N and extending to D are only possible in Mandarin, and spans starting at Cl and extending to D are only possible in Cantonese. Other spans are the same in both languages.

4. A head movement account and its limitations

Simpson (2005) and Wu & Bodomo (2009) propose a head movement account of the variety of interpretational restrictions on different noun phrase configurations in Cantonese, generally following the insights of Longobardi (1994). Cheng & Sybesma (1999) also include head movement as part of their explanation for the differences between Mandarin and Cantonese. The proposals differ in their assumptions about the locus of definiteness, and about universals of syntactic structure. However, one fact about each head movement analysis is that it makes the prediction that a definite bare noun which is modified by an adjective should precede that adjective, which is contrary to what we actually see in Mandarin.

Cheng & Sybesma argue that the Cl head in Mandarin and Cantonese plays the (semantic) role that D does in English. They state that there is a condition on interpreting Cl as an iota operator, and that this condition is simply that the Cl position be filled.¹⁷ In Cantonese, the classifier is overt, and thus the Cl head is filled. In Mandarin, N moves to the Cl projection in the case of a definite bare noun. Given that the canonical order of elements in the DP in Mandarin is Num > Cl > Adj > N, N movement to Cl would give rise to the order N > Adj when N is definite, which does not hold for Mandarin (the order N > Adj is always unacceptable).

Simpson (2005) and Wu & Bodomo (2009) posit a D projection as the locus of definiteness, and take it to be Cl moving to D that gives rise to a definite interpretation in Cantonese. They do not discuss Mandarin, but one could imagine extending this analysis by arguing that in Mandarin, N moves to D for a definite interpretation. In that case, we have the same problem that the Cheng & Sybesma account faces, namely that the wrong order is predicted when a definite bare noun is modified by an adjective.

An alternative would be to claim that it is the Cl head which moves to D even in Mandarin, and that Cl is null in the case of bare nouns. However, this would make it necessary to make some quite unusual stipulations about the phonological content of Cl in different contexts. Without going into details, one would have to claim that in Mandarin, a definite interpretation (with movement of Cl) would force the Cl to be phonologically null

¹⁷ On this point my account and Cheng & Sybesma's are very similar, in that both accounts posit a morphological licensing condition on the head which hosts the iota operator.

(only bare nouns are definite), and an indefinite interpretation (with no movement of Cl) would mean that Cl could be either null or overt (bare nouns and Cl-N phrases can both be indefinite). A null head moving to a null head to license a particular feature on the higher head seems unsatisfying as a solution to the problem.¹⁸

Overall, the various alternatives of the head movement account mean that we either get the word order wrong when adjectives are taken into consideration, or we have to make some very unwelcome stipulations about what kind of phonological content is spelled out in the Cl head.

As discussed above (section 3.4), the low spell-out of the span in (14) means that the adjective precedes the (bare) noun even where we have a definite interpretation. It is this characteristic of the spanning approach which makes it superior to a head-movement approach in this case; head movement forces the head up the tree, but the ordering of elements suggest that it is pronounced low. A non-movement account thus captures the data more straightforwardly.

The spanning approach gets the word order right, and leads to generalizations which do not require the same kind of stipulations that we need to force the word order to work in the head-movement approach.

5. Crosslinguistic evidence

Given the different possible licensing conditions on D presented in section 3.3 (8), one would expect that this opens up a number of potential options that languages have available to them, and that therefore different languages would make use of different sets of the options available. Logically there are 8 possibilities, summarized below:

- (20) a. Article only
- b. Span spell-out only
- c. Filled Spec only
- d. Article & Span
- e. Filled Spec & Span
- f. Article & Filled Spec
- g. Article & Filled Spec & Span
- h. No definite interpretation

If we take possessive DPs to involve a phrasal possessor in spec D (which licenses a definite reading), then it is likely that no languages make use of the (a) or (b) options, unless possession is realized in some other way (Romance languages might be such a case). Below I give the examples of Bangla as a language that looks like option (c), and Nuosu Yi as a language that looks like option (d). I have claimed in this paper that MC and Cantonese are languages which exemplify option (e). (f) looks not unlike English, again if we take possession to involve a filled spec D position which licenses a definite D head. I do not know of a language matching option (g), but hope that further cross-linguistic investigation would

¹⁸ Another alternative would have the movement of N to D/Cl to be LF head movement. This solution is unsatisfying, particularly given that head movement is known to generally not give rise to any semantic effects (Chomsky 2001, Adger 2013; see Lechner (2005) and Roberts (2010) for arguments that there are interpretive effects of head-movement).

shed light on the possibility of its existence. I expect that option (h) is not a possible language, on the assumption that something like definiteness (that is, a D projection) and the constraint on its being licensed are given *a priori*, and that the task of the child acquiring a language is to try to map parts of the primary linguistic data onto potential licensors of functional elements, including D. In acquisition, the child has to work out not only what is flagging D, but also which types of flagging are available.

I will now briefly present two interesting cases out of the possibilities that I presented above, and also illustrate that the analysis of MC and Cantonese provides a simple explanation for the behaviour of a definite article in a classifier language, Nuosu Yi.

5.1. Span spell-out and article spell-out

An interesting case is Nuosu Yi, a Tibeto-Burman classifier language which has a definite article *su*, and which has bare nouns that can be interpreted as either definite or indefinite. I present here a summary of the possible noun phrase configurations and their related interpretations (summarized from data presented in Jiang 2013 and Gerner 2014):

Table 5: Noun phrase interpretation in Nuosu Yi

<i>Noun Phrase Config</i>	<i>Interpretation</i>
N Cl	[--def] (singular)
N Cl <i>su</i>	[+def] (singular)
N	[±def] (number neutral)
N <i>su</i>	*

Given that bare nouns can be interpreted as definite, we assume (as with MC above) that there is a span available from N to D, licensing a definite interpretation. The language also has a definite article *su*, which forces a definite interpretation when it combines with a N-Cl sequence. What is relevant to the discussion here is that a noun immediately followed by the definite article is ungrammatical. This fact is neatly explained by the spanning account given above: a bare noun with a definite interpretation is the spell-out of a span which includes the D head, and so that head could not spell out as a separate article (assuming that one is forced to spell out the largest possible span available). There is only one D head; either it is included in the span and spelled out with it, or it is spelled out separately as an article.

5.2. Filled spec only

Bangla is potentially an example of a language which only makes use of the filled spec licensing option. Summarizing data from Bhattacharya (1993) and Dayal (2012), we have the following paradigm:

Table 6: Noun phrase interpretation in Bangla

<i>Noun Phrase Config</i>	<i>Interpretation</i>
Numeral CI [N]	[-- def]
[N] Numeral CI	[+ def]
Numeral CI [A N]	[-- def]
[A N] Numeral CI	[+ def]
CI N	*
N CI	[+ def]

Bhattacharya (1993) and Dayal (2012) have both argued that in its non-canonical position at the left edge of the DP, [N] or the constituent [A N] have undergone phrasal movement. In each case, this phrasal movement gives rise to a definite interpretation. Bare nouns in Bangla cannot have a definite interpretation (Dayal 2012), and so presumably there is no span from N to D available, suggesting that filled-spec is the only licensing possibility available in Bangla.

6. Conclusion

I have shown that differences in possible interpretations for different noun phrase configurations in MC and Cantonese, and the interpretational effects associated with a high adjective can be explained straightforwardly by an account which views the syntax-morphology interface to be mediated through the spell-out of spans. I have also shown that such an approach is superior to one which attempts to explain the same effects through head-movement of N (or CI) to D.

The above argument leads to predictions about possible parametric variation, and discussion of Nuosu Yi and Bengali provides preliminary evidence that those predictions may be borne out (following further cross-linguistic investigation). Another consequence of the discussion is that a parsimonious account of the facts can be achieved through the positing of a D functional projection even in languages which have no overt articles. This suggests that the recent attempts to argue that the DP hypothesis does not hold for argument nominals in languages like Mandarin (see, e.g., Bošković 2008) might require rethinking; here we have another case where surface dissimilarity obscures, but does not preclude, underlying parallelisms.

References

- Abels, K. & Muriungi, P. (2008). The focus marker in Kĩharaka: Syntax and semantics. *Lingua* 118, 687-731.
- Adger, D. (2013). *A syntax of substance*. Cambridge MA: MIT Press.
- Bhattacharya, T. (1993). *The structure of the Bangla DP*. Ph.D. dissertation, University College London.
- Borer, H. (2005). *In name only: Structuring sense volume 1*. Oxford: Oxford University Press.
- Borer, H. (2013). *Taking form: Structuring sense volume 3*. Oxford: Oxford University Press.
- Bošković, Ž. (2008). What will you have, DP or NP? *Proceedings of NELS* 37.
- Brody, M. (2000). Mirror theory: Syntactic representation in perfect syntax. *Linguistic Inquiry* 31, 29-56.
- Cheng, L. & Sybesma, R. (1999). Bare and not-so-bare nouns and the structure of NP. *Linguistic Inquiry* 30, 509-542.

- Cheng, L. & Sybesma, R. (2005). Classifiers in four varieties of Chinese. In Cinque, G. & Kayne, R. (eds.), 259-292.
- Cinque, G. (1994). On the evidence for partial N-movement in the Romance DP. In Cinque, G., Koster, J., Pollock, J.-Y., Rizzi, L., and Zanuttini, R. (eds.) *Paths towards Universal Grammar: Studies in honor of Richard S. Kayne*, 85-110. Washington, DC: Georgetown University Press.
- Cinque, G. & Kayne, R. (eds.) (2005). *Handbook of comparative syntax*. New York: Oxford University Press.
- Chomsky, N. (2001). Derivation by Phase. In Kenstowicz, M. (ed.), *Ken Hale: A life in language*, 1-52. Cambridge MA: MIT Press.
- Dayal, V. (2012). Bangla classifiers: Mediating between kinds and objects. *Italian Journal of Linguistics* 24.
- Dékány, E. (2011). *A profile of the Hungarian DP*. Ph.D. dissertation, University of Tromsø.
- Diesing, M. (1992). *Indefinites*. Cambridge MA: MIT Press.
- Elbourne, P. (2013). *Definite descriptions*. Oxford: Oxford University Press.
- Gerner, M. (2013). *A grammar of Nuosu*. De Gruyter Mouton.
- Grimshaw, J. (2005). Extended Projection. In Grimshaw, J. (ed.) *Words and structure*. 1-74. Stanford, CA: CSLI Publications.
- Harbour, D. (2014). Paucity, abundance and the theory of number. *Linguistic Inquiry* 90, 185-229.
- Heim, I. (1982). *The semantics of definite and indefinite noun phrases*. Ph.D. dissertation, University of Massachusetts.
- Huang, J. (1982). Logical relations in Chinese and the theory of grammar. Ph.D. dissertation, MIT.
- Huang, J. & Li, A. & Li, Y. (2009). *The syntax of Chinese*. Cambridge: Cambridge University Press.
- Jenks, P. (2012). Definite spans and blocking in classifier languages. Ms. University of California, Berkeley.
- Jiang, L. (2012). *Nominal arguments and language variation*. Ph.D. dissertation, Harvard University.
- Lin, J. & Zhang, N. (2006). The syntax of non-referential *ta* 'it' in Mandarin Chinese. *Language and Linguistics* 7, 799-824.
- Lu, B. (1998). *Left-right asymmetries of word order variation: A functional explanation*. Ph.D. dissertation, University of Southern California.
- Simpson, A. (2005). Classifiers and DP structure in Southeast Asia. In Cinque, G. & Kayne, R. (eds.), 806-838.
- Sio, J. (2006). *Modification and reference in the Chinese nominal*. Ph.D. dissertation, The Leiden University Centre for Linguistics.
- Svenonius, P. (2012). *Spanning*. Ms. CASTLE, University of Tromsø.
- Sybesma, R. (1999). *The Mandarin VP*. Dordrecht: Kluwer Academic Publishers.
- Wu, Y. & Bodomo, A. Classifiers ≠ determiners. *Linguistic Inquiry* 40, 487-503.
- Yang, H. (2005). Plurality and modification in Mandarin nominal phrases. Ph.D. dissertation, University of Texas at Austin.
- Zhang, N. (2015). Nominal internal phrasal movement in Mandarin Chinese. *The Linguistic Review* 32 (2).

*David Hall
Department of Linguistics (SLLF)
Queen Mary University of London
Mile End Road
London E1 4NS*

d.t.hall@qmul.ac.uk

CHILDREN'S INTERPRETATION OF JAPANESE PARTICLES IN COMPLEX SENTENCES

KEIKO HATA

(University of Hawai'i at Mānoa)

Abstract

This paper discusses how monolingual children speaking Japanese as their first language interpret the particles, *WA* and *GA*, in complex sentences. Previous studies (Nakaiwa et al. 1995, Uchida, et al. 1995, Nariyama 2002) reported adult Japanese-speakers' tendency of interpreting *WA* marking as a coreference with the elided subject while *GA* marking as a different subject. The current study examined children speaking L1 Japanese in a picture-selection task. Results show that children distinguished the particles similarly to adults in most conditions, while performing differently when *WA* appeared in the middle of a sentence.

1. Introduction

This paper discusses how children whose native language (L1) is Japanese interpret the particles *WA* and *GA* in complex sentences. *WA* and *GA* are particles which are attached to noun phrases (NPs). Although [NP-*WA*] and [NP-*GA*] can be placed in the subject or object position in a sentence, the focus in this study is on those in the subject position. Functional distinctions of each particle are ambiguous when given out of context. Japanese linguists have compared these particles and provided various types of usages and functions, such as contrastive/exhaustive, old information/new information, topic/ subject, etc. (Kuno 1973, Hinds, Maynard, & Iwasaki 1987, Noda 1996 among others). However, these definitions seem to be determined or interpreted subjectively, varying from person to person and from context to context.

Japanese as being as null-subject language allows elided subjects at roughly 70% in conversation and 50% in written narrative texts (Hinds 1983, Mizutani 1985, Nariyama 2000). How Japanese speakers determine the referential identity of elided subjects is explained as depending on contextual cues, structures (ex. active vs. passive), verbs (ex. transitive vs. intransitive), honorific expressions, etc. However, it appears that resolution for elided subjects in complex sentences are rather systematized and that adult Japanese speakers depend on the systems by distinguishing *WA* and *GA* consciously or unconsciously since it is attested that within a complex sentences *WA* indicates the identity of an elided subject as a coreference with the *WA*-marked subject and that *GA* more likely indicates the identity of an ellipsis as a different subject from the *GA*-marked subject (Narikawa et al. 1995, Uchida et al. 1995, Nariyama 2002). Although this *WA/GA* distinction in complex sentences seems to be a very common phenomenon in adult Japanese, how L1 Japanese-speaking children interpret each particle has yet to be investigated. With that said, the aim of the current study is to test whether L1 Japanese-speaking children use the aforementioned system to distinguish between *WA* and *GA* to resolve elided subjects in complex sentences, particularly those with two clauses: adverbial and matrix, in the same manner as adult L1 speakers do.

2. Literature Review

Previous studies on *WA/GA* distinction in complex sentences show that adult L1 speakers of Japanese make distinctions between these particles to identify elided subjects.

These studies reported that an elided subject is coreferential with a WA-marked subject in a complex sentence, and that GA tends to invite interpretation of switch-reference.

2.1. Examination of Translation Texts

Nakaiwa et al. (1995) examined translation from Japanese to English by L1 Japanese speakers. Out of 3781 complex sentences, which have zero pronouns making intrasentential and extrasentential anaphoric references, there were 515 zero pronouns in the locations of subject, direct object, indirect object, and others. 124 of them were in the subject location and their antecedent in the same sentence. Nakaiwa et al. found that 109 of them (88%) were translated as having subjects coreferential with WA-marked subjects, while only 8 of them (6%) were coreferential with GA-marked subjects. This finding indicates that 88% of coreferences were signaled by WA-marked antecedents and thus WA can be the indicator binding two subjects in a sentence as the same entity at significant frequency, while GA rarely triggers a coreferencing reading.

2.2. Sentence-Completion Task

Uchida et al. (1995) conducted a sentence completion test, targeting 67 adult native speakers of Japanese, to see how WA-, or GA-marked subjects in a subordinate clause¹ influence the reader's choice of subject in the following clause. Participants were asked to create main clauses which may follow given subordinate clauses. These subordinate clauses were presented in two different conditions: a minimal pair with the only difference being the subject markings of WA and GA. The responses show that the native speakers interpreted all WA-subject as an antecedent of the subject in the main clause (100%); on the other hand, 56% of GA was interpreted as marking different subjects. Although the percentage in the GA case appears to be too low for the particle to be claimed as the marker of a different subject from an ellipsis, this rather low percentage can be explained by experimental sentences being presented out of context.

2.3. Switch-Reference Systems in Japanese

Nariyama (2002) examined written narrative texts and argued that the interaction of WA and GA in complex sentences has an analogous property of switch-reference systems which determine the referential identity of elided subject. That is, WA signals Same Subject; GA signals Different Subject. She analyzed the WA/GA distinction in complex sentences as being similar to the function of switch-reference systems observed in Mojave language. Comparing the WA/GA distinction to the Mojave switch-reference systems, Nariyama (2002) proposed that WA, just like the marker *-k* in Mojave, signals Same-Subject (SS) which denotes that the subject of the marked, or subordinate, clause is the same subject as the subject in the main clause; on the other hand, GA, just like the marker *-m* in Mojave, signals Different-Subject (DS).

To sum up, as these studies show, there is a strong tendency that L1 Japanese-speaking adults use given particles to judge the identity of elided subjects in complex sentences. However, it seems that existing studies on the WA/GA distinction for elided subject resolution in complex sentences have examined written texts and been conducted with L1 Japanese-speaking adults only. In other words, no previous studies apparently address how L1 Japanese-speaking children resolve identifying elided subject in complex sentences.

¹ In Japanese, it is canonical that adverbial clauses are followed by the matrix clauses (Kuno, 1978).

3. Present Study

3.1. Method

I investigated the issue above by collecting data from children speaking L1 Japanese in a picture-selection task. The focuses of analysis were on whether children make distinctions between WA and GA in complex sentences, and if they do, how similarly or differently these distinctions are made in comparison to adults, in order to answer the research questions as follows:

- (1) Do children make distinction between WA and GA to resolve elided subject in complex sentences?
- (2) If they do, how similarly or differently do children interpret functions of WA/GA in complex sentences?

3.2. Participants

To investigate children's reactions to WA and GA in complex sentences, 5 monolingual Japanese children (Mean age 6;11, range 6;8–7;3) and 14 adult native speaker controls (Mean age 29, range 26–42) participated in a picture-selection experiment.

3.3. Materials

The experimental materials used in the picture-matching task were 5 sets of 15 sentences (5 test items and 10 fillers in each set). A total of 25 test items were prepared; 10 were composed of a *toki* 'when' clause and matrix clause, and 10 were composed of a *kara* 'because' clause and matrix clause. Among Japanese conjunctures, *toki* and *kara* were chosen due to their semantic features which can denote the simultaneousness of two different events expressed in subordinate and matrix clauses separately, unlike *maeni* 'before' or *atode* 'after.' Furthermore, conjunctions which represent 'SS' only, such as *nagara* 'while (V1)ing / at the same time S + V2.' as in *Hanako-wa terebi-o mi-nagara, juusu-o nonda* 'Hanako, while watching TV, drank juice' and *te* 'X + V1 and then V2,' as in *Hanako-wa terebi-o mi-te benkyoo-shita* 'Hanako watched TV and then studied,' were disregarded in this study. (See Appendix 1 for sample scripts, including context sentence and test item). These sentences were audio-recorded and set to be played by clicking a speaker icon on a Power Point slide. On slides, corresponding pictures to the sentences, each of which depicts two different scenarios with one of the characters as a narrator (See Appendix 2 for sample pictures).

3.4. Procedures

(1) On a power-point slide, two contrastive pictures appear with a click. Participants were asked to describe each of the pictures. The purpose of this step is to provide the participants with an opportunity to examine the pictures and to make sure they know what differentiates them as to who did what. (2) On the second click, the narrator, one of the characters in given pictures, appears so that participants know who is the speaker of a forthcoming statement. (3) By clicking the audio speaker icon, which is placed in the center of the screen, participants hear the context and the narrator's statement. (4) After listening to the narrator, participants were directed to choose one of two pictures to indicate their interpretation of what the narrator has just said.

3.5. Results

Results indicate that children identified elided subjects in complex sentences in similar manners as adults did, in 4 out of 5 conditions. In the condition where a WA-marked subject appears in the middle of a sentence, children performed very differently from adults, shown in Tables 1 and 2 below:

Table 1. Adults' interpretation of identifying elided subject (n=14).

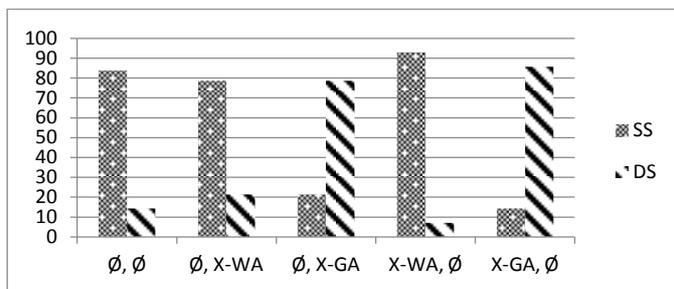
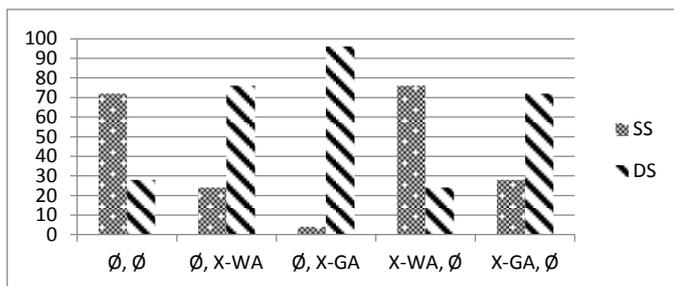


Table 2. Children's interpretation of identifying elided subject (n=5).



These findings may suggest that children by the age of 6 are sensitive to the WA/GA distinction in complex sentences, although the data collected is too small for generalization of such a linguistic phenomenon. Nonetheless, it can be an interesting observation that children at this age have tendency to interpret GA as the indicator of Different Subject (DS), wherever it is located, which is adult-like. When WA came in the front of the sentence, children more likely interpreted the WA-marked subject to be coreferential with an elided subject in the main sentence, or Same Subject (SS). This again turns out to be similar to the adult interpretation. In the case with null subjects in both clauses, they were more likely interpreted as coreferencing with each other (SS) identified as the narrator in given contexts (ex. ‘Frog’ talking about the incident when he had an argument with ‘Snail’).

What may be most interesting in the data is the Condition 2 where WA appears in the middle of a sentence. While adults showed their tendency of a Same Subject reading in this condition, 76% of the time children took a DS reading although they seem to have understanding that WA and GA are different markers functioning differently.

4. Discussion

The results have revealed differences between the way children and adults interpret sentences in Condition 2. Specifically, the findings indicate that whereas adult native speakers of Japanese are rather influenced by the particles given in a sentence, rather than their locations, as attested from the previous studies, children may read given sentences linearly as they hear. In other words, children at the age of 6 distinguish the particle, WA or GA, marking

the NP in the very first position to use it as a tool to determine the identity of elided subjects following given conjunctions, *toki* or *kara*, as SS or DS (Conditions 4 and 5). These children also know that null subjects in the beginning of a sentence can be a coreference with the narrator (Conditions 1, 2, and 3), that the narrator keeps its identity with the other elided subject when no WA/GA-marked NP appears after a conjunctive (Condition 1), and that GA in the matrix subject position, following a clause with an elided subject, indicates someone else other than the narrator (Condition 3). Perhaps, in child Japanese, the particles, whether WA or GA, are indicators of DS when they appear in the middle of a sentence, having no functional differences from each other.

This interpretive difference of WA/GA placed in the middle of a sentence observed in child Japanese and adult Japanese may be attributed to differences in their linguistic processing skills which children at this age may not have yet developed. Among all the given conditions, Condition 2 represents relatively complicated structures where the WA-marked NP, although appearing to belong to the second clause, is actually binds the whole sentence to corefer with the elided subject in the first clause. For children who process sentences linearly as they hear, Condition 2 is challenging in that its structure requires reanalysis to trace back to the elided subject in the first clause to revise interpretation of the identity of the ellipsis from as the narrator to instead being someone else that is marked with WA. This is a complex processing and perhaps it is too complicated for 6-year old children to perform such reanalysis.

5. Conclusion

This study shows that children by the age of 6 understand the functions of WA and GA in complex sentences when the given structures are straightforward, whereby they make the WA/GA distinction for elided subject resolution mostly in the same manner as adults do. However, more data of both child and adult Japanese should be collected to generalize this particular linguistic phenomenon. Also it is necessary to refine stimuli sentences and their corresponding pictures to collect more precise data. Data from a younger population may provide further insight into innateness of child language regarding WA/GA distinctions.

Appendix 1

Context: Frog and Snail had a fight the other day. Frog is talking about that day.

<Condition 1>

\emptyset *okotta* *kara*, \emptyset *kaetta-nda*.
 \emptyset was-angry because \emptyset went-home
 ‘Because \emptyset was angry, \emptyset went home.’

<Condition 2>

\emptyset *okotta* *kara*, *Katatumuri-wa* *kaetta-nda*.
 \emptyset was-angry because Snail-WA went-home
 ‘Because \emptyset was angry, Snail went home.’

<Condition 3>

\emptyset *okotta* *kara*, *Katatumuri-ga* *kaetta-nda*.
 \emptyset was-angry because Snail-GA went-home
 ‘Because \emptyset was angry, Snail went home.’

<Condition 4>

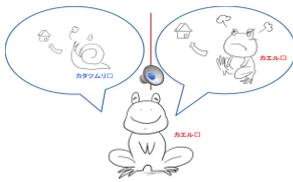
Katatumuri-wa okotta kara, ø kaetta-nda.
Snail-WA was-angry because ø went-home
'Because Snail was angry, ø went home.'

<Condition 5>

Katatumuri-ga okotta kara, ø kaetta-nda.
Snail-GA was-angry because ø went home
'Because Snail was angry, ø went home.'

Appendix 2

For Condition 1:



For Conditions 2 & 3:



For Conditions 4 & 5:



References

- Hinds, J., Maynard, S., and Iwasaki, S. (1987). *Perspectives on Topicalization: The Case of Japanese wa*. Amsterdam: John Benjamins.
- Kuno, S. (1973). *The Structure of the Japanese Language*. Tokyo: Taishukan.
- Kuno, S. (1978). *Danwa no Bunpo (Discourse grammar)*. Tokyo: Taishukan.
- Nakaiwa, H., Shirai, S., Ikehara, S., & Kawaoka, T. (1995). Extrasentential resolution of Japanese zero pronouns using semantic and pragmatic constraints. AAAI 95 Spring Symposium, 99–105.
- Nariyama, S. (2000). *Referent Identification for ellipted arguments in Japanese*. PhD dissertation, University of Melbourne.

- Nariyama, S. (2002). The WA/GA distinction and switch-reference in Japanese complex sentences. *Studies in Language* 26(2), 369–431. Amsterdam: John Benjamins.
- Noda, H. (1986). *Fukubun ni okeru 'wa' to 'ga' no Kakari-kata. (Biding of 'wa' and 'ga' in complex sentences)*. Tokyo: Meiji Shoin.
- Uchida, A., Ikegami, M., Ono, S., Oshima, Y., & Nagatomi, K. (1995). Yosoku Bunpoo Kenkyuu: GA to WA no Yosoku-kinoo nitsuite (Study of expectancy grammar: Predictive function of GA and WA). *Gengo bunka to nihongo kyouiku (Culture of language and Japanese education)*: 134–153. Nihon Gengo-bunka-gaku Kenkyujo

Keiko Hata
Department of East Asian Languages and Literatures
University of Hawai‘i at Mānoa
1890 East-West Road
Honolulu, Hawai‘i 96822
United States of America

keiko3@hawaii.edu

GENERIC PRONOUNS AND PHI-FEATURES: EVIDENCE FROM THAI

ANDERS HOLMBERG

Newcastle University and University of Cambridge

ON-USA PHIMSAWAT

Burapha University

Abstract

In this paper, we investigate the triggers and the distribution of generic pronouns in regard to the theory of pronouns, ϕ -features, reference, and pro-drop. We survey the empirical evidence for different ‘pro-drop types’ focusing on generic pronominal expressions in a set of languages, namely radical pro-drop languages, consistent pro-drop languages, and partial pro-drop languages. In particular we are concerned with the properties that are responsible for the difference between the inclusive, quasi-inclusive, and exclusive generic pronouns, and whether or not the generic reference includes reference to non-humans along with humans. We show that the inclusive pronoun in Thai has unrestricted reference, including humans and (potentially) non-humans. The explanation is that this pronoun has no phi-features. This also explains why it is null: There are no phi-features to spell out. We argue that the quasi-inclusive and exclusive generic pronouns include a silent noun ‘people’.

1. Introduction¹

There are three types of generic pronouns, inclusive, quasi-inclusive, and exclusive, as exemplified in (1a,b,c), respectively.

- (1) a. **One** should always be in love. That’s the reason **one** should never marry.
(Oscar Wilde)
- b. **We** like smoked fish in Finland.
- c. **They** speak lots of different languages in India.

The inclusive pronoun is so called because it refers to people in general including the speaker and the addressee, while the quasi-inclusive pronoun refers to people in general including the speaker but not the addressee, and the exclusive pronoun refers to people in general, in some domain, excluding the speaker and the addressee. There is cross-linguistic variation regarding generic pronouns, including when they can be, or must be null (Holmberg 2005, 2010b, Sigurðsson and Egerland 2009, Phimsawat 2011). The purpose of the paper is to contribute to the description and explanation of this variation in terms of a theory of pronouns, ϕ -features, and reference.

¹ Anders Holmberg’s research on this paper is funded by the European Research Council Advanced Grant No. 269752 ‘Rethinking Comparative Syntax’ (ReCoS). Thanks to the other members of the ReCoS group, Ian Roberts, Theresa Biberauer, Michelle Sheehan, Jenneke van der Wal, Andras Barány, Georg Hahn, for discussing these issues with us. We would also like to thank especially the following people for their contribution to this paper: Peter Arkadiev, Seiki Ayano, Maia Duguine, Mara Frascarelli, Yujia Han, Shin-Sook Kim, Tawee Kueakoolkiat, Makiko Mukai, Halldor Sigurðsson, Ur Shlonsky, Salinee Somtopcharoenkul, Harold Thampoe, Phan Trang, and Hofa Meng Jung Wu.

A key question is, what features do generic pronouns have? And insofar as we can determine this, what does this tell us about ϕ -features and the nature of pronouns more generally? We will start by a survey of generic pronominal expressions in a set of languages representing different ‘pro-drop types’. In this paper we will focus mainly on Thai, a representative of the radical pro-drop type, with subject and object pro-drop without involvement of agreement.

We will show that the syntactic basis for the interpretation of the generic pronouns, though similar in the case of the exclusive and quasi-inclusive pronouns, is quite different in the case of the inclusive pronoun.

2. Generic pronouns in different types of pro-drop languages

In Thai, a radical pro-drop language, the generic pronouns are realised as follows:

- (2) a. *díawníi ɲaan hǎa yâak mâak thâa Ø mây còb trii.* [Thai]
 nowadays job seek difficult very if NEG finish BA
 ‘Nowadays to seek a job is difficult if one hasn’t finished a BA.’
- b. *raw kin cee nay duan tùlaakhom.*
 we have vegetarian food in month October
 ‘We have vegetarian food in October.’
- c. *bon kò ní sùanyài (khǎw) plùuk chaa khǎay.*
 on island DEM mostly they grow tea sell
 ‘On this island they grow and sell tea.’

The inclusive pronoun in (2a) is null, and can only be null; there is no overt inclusive pronoun in Thai. The quasi-inclusive pronoun *raw* in (2b) is overt, and has to be, if the sentence is uttered out of the blue. The pronoun in (2c), on the other hand, can be null or overt, and still be interpreted as generic, even if the sentence is uttered out of the blue,

Consider Italian, an example of a consistent pro-drop language, in the terminology of Holmberg (2005) and Biberauer et al. (2010).

- (3) a. *Si lavora sempre troppo.* [Italian]
 SI work.3SG always too-much
 ‘One always works too hard.’
- b. *Secondo il primo ministro, (noi) dobbiamo essere p ú produttivi.*
 According the prime minister we should.1PL be more productive
 ‘According to the PM, we need to be more productive.’
- c. *Ø parlano molte lingue in India.*
 speak.3PL many languages in India
 ‘They speak many languages in India.’

As discussed in Holmberg (2005) and Biberauer et al. (2010) the inclusive generic pronoun cannot be null in Italian and other consistent pro-drop languages. More precisely, it cannot be realised as a null, 3SG pronoun, a null ‘one’, in an ordinary active sentence. Sentence (6) can only be interpreted as shown, as having a referential 3SG subject.

- (4) Lavora sempre troppo. [Italian]
 work.3SG always too-much
 ‘He/she always works too hard.’

In Italian and most other Romance languages, the inclusive generic sentence employs the reflexive clitic pronoun *si/se*. The structure of sentences employing this pronoun is controversial; see Cinque (1988), Dobrovie-Sorin (1998), D’Alessandro (2007). The pronoun may be a realisation of the generic subject. Alternatively, it is a voice marker licensing a null, impersonal pronoun. In either case, it holds that ‘something special’ is required because the inclusive generic pronoun cannot be realised as a null 3SG pronoun in an ordinary active, finite sentence. The Slavic consistent pro-drop languages also make use of a reflexive pronoun in inclusive generic sentences (Krzek 2013a,b). Other consistent pro-drop languages resort to an overt indefinite pronoun, or passive voice, or they fall back on the 2^{SG} inclusive generic; see Holmberg (2010b).

The quasi-inclusive pronoun, on the other hand, is often null, but can be overt,² while the exclusive subject pronoun has to be null; see (3b,c).

Now consider Finnish, a partial pro-drop language, in the terminology of Holmberg (2005) and Biberauer et al. (2010).

- (5) a. T ä s s ä istuu Ø mukavasti. [Finnish]
 here sits comfortably
 ‘One can sit comfortably here.’
- b. (Me) sy ö m m e Suomessa paljon savukalaa.
 we eat.1PL Finland.INE much smoked.fish
 ‘We eat a lot of smoked fish in Finland’
- c. Intiassa puhutaan Ø monta eri kielt ä
 India.INE speak.IMPL many different language
 ‘They speak many different languages in India.’

Characteristic of partial pro-drop languages is that they allow pro-drop but under more restricted circumstances than consistent or radical pro-drop languages. In Finnish, subject pro-drop is generally optional with 1st and 2nd person pronouns, but with 3rd person referential pronouns it is only possible in embedded position under control by a subject in a higher clause; see Holmberg, Nayudu & Sheehan (2009), Holmberg (2010a), Holmberg & Sheehan (2010). As shown in (5a), the inclusive generic pronoun is null, though, and has to be (the position of the null pronoun is because the finite verb undergoes raising to T). As shown in (5b), the quasi-inclusive pronoun can be optionally null. The exclusive generic construction is generally expressed in Finnish with an impersonal verb form, also identified as a passive (Blevins 2003, Manninen and Nelson 2004). It is controversial whether there is a null subject in the construction at all.³

The upshot is that the quasi-inclusive generic pronoun behaves in a similar manner in the three languages. We will see below that it can be null in Thai, too, given the right context.

² The following example was found on the internet. Thanks to Mara Frascarelli for discussing this point with us.

(i) praticamente secondo loro **noi** dovremmo credere a Bersani, Berlusconi, o Monti ma..."
 in other words, in their opinion we should believe Bersani, Berlusconi or Monti but..."

³ An interesting complication in Finnish is that the 1PL, referential as well as generic, is colloquially expressed by the same impersonal form as in (5c). In that case, the pronoun has to be overt. We will ignore this complication here.

Even putting the Finnish case aside, there is variation regarding the exclusive generic pronoun, though: optionally null in Thai, but obligatorily null in Italian. And especially with regard to the inclusive pronoun there is some striking variation: null in Thai and Finnish, but not in Italian. As we shall see below, there is also an interesting difference between the null generic pronoun in Thai and that in Finnish. Thus it is different in the three languages, the differences correlating with the type of pro-drop, as we will demonstrate.

These properties of the generic pronouns, we will argue, are not accidental, language-particular facts, but follow, for the most part, from a theory of pronominal reference and pronominal features.

3. The inclusive generic pronoun

We will begin by recounting the explanation in Holmberg (2005, 2010) for why Italian and other consistent pro-drop languages do not have a 3SG generic pronoun, while Finnish and the other partial pro-drop languages do. The difference is to do with the ϕ -features of T, i.e. with agreement: the consistent pro-drop languages include a definiteness feature as part of the ϕ -feature set of T, the partial pro-drop languages don't. In this respect, the consistent pro-drop languages have richer agreement than the partial pro-drop ones. The effect when T has a definiteness feature is that when a definite 3SG pronoun enters an agreement relation with T, all its features are copied by T, and as a result, the pronoun is deleted/ not spelled out, being a copy of the features of T. An indefinite or generic 3^{SG} pronoun cannot be null; it will not be a copy of the features of T. In Italian, as mentioned, the inclusive generic pronoun is spelled out as *si*. In the partial pro-drop languages, T does not include a definite feature. This means that the features of a definite 3rd person subject pronoun will not all be represented in T. Consequently the definite 3rd person pronoun cannot be deleted (except if it is controlled by a nominal argument in a higher clause; see Holmberg, Nayudu and Sheehan 2009, Holmberg and Sheehan 2010). The inclusive generic pronoun, according to Holmberg (2005, 2010), has 3SG features but no D-layer. When it enters an agreement relation with T, all its features will be represented by the D-less T, and consequently the pronoun will be deleted, being a copy of the features of T.

What features do generic pronouns have? What is it that makes them generic, not referential? As for the quasi-inclusive generic pronoun 'we' and the exclusive 'they', it seems straightforward enough that they have 1PL and 3PL features, respectively. Their null counterparts in Italian also have these features, judging from the agreement on the finite verb. We repeat the relevant examples with null subjects.

- (6) a. Secondo il primo ministro, dobbiamo essere p _i produttivi. [Italian]
 according the prime minister should.1PL be more productive
 'According to the PM, we need to be more productive.'
- b. Parlano molte lingue in India.
 speak.3PL many languages in India
 'They speak many languages in India.'

Still, the fact that they have generic reference rather than specific suggests that there is some featural difference between them and specific 1PL and 3PL pronouns. We will put the quasi-inclusive and exclusive pronouns aside for the time being, and instead focus on the inclusive generic pronoun. In English, generic *one* is 3SG. In Finnish, too, the agreement on the verb signals that the generic subject pronoun is 3SG.

(7) a. One always works too hard.

b. T ä n n e t u l e e m i e l l ä ä n . [Finnish]
here come.PRS.3SG with.pleasure
'It's nice to come here.'

While this is uncontroversial in the case of English, it is not as straightforward in Finnish. There is a possibility that the 3SG marking on the verb is not due to agreement but is a default marking, in the absence of anything to agree with. But as discussed in Holmberg (2005, 2010b), there are some clear indications that the 3SG on the verb is due to agreement with the null subject. We will not repeat that discussion here. Also in other languages with an overt inclusive generic pronoun, it is typically 3SG. There are also languages where it is 3PL, among them Hebrew and, with some qualification, Russian, two partial pro-drop languages (Barbosa, to appear).⁴

There are also many languages where the inclusive generic meaning can be expressed by a 2SG pronoun, overt or null, with 2SG agreement, if the language has agreement. We put this generic pronoun aside in this paper, though (see Gruber 2013).

What about Thai? Thai is a radical pro-drop language with no agreement and no pronounced inclusive generic pronoun.

(8) d i a w n i i n a a n h ä a y ä a k m ä a k t h ä a Ø m ä y c ò b t r i i [Thai]
nowadays job seek difficult very if NEG finish BA
'It's difficult to seek a job nowadays if one hasn't finished a BA.'

Consider the semantics of the inclusive generic pronoun. As mentioned, the meaning is 'people in general including me and you'. Since the speaker and the addressee are included in the reference of the pronoun, there is actually no semantic motivation for the 3rd person feature that the generic pronoun has in English, Finnish and many other languages. The reference includes everybody, speaker, addressee and everybody else: it is unrestricted. We take this to mean that the inclusive generic pronoun has no ϕ -features. This follows given that what ϕ -features do is restrict the reference of a pronoun (or nominal expression more generally), to only the speaker (1SG), or the speaker and his/her associates (1PL), or the addressee (2SG), or a female person who is not the speaker or the addressee (3SG.F), etc.

This suggests that the 3SG feature is prevalent as marking of the inclusive generic pronoun in some languages because it is the minimal ϕ -feature specification. In some theories of pronominal features 3SG is a minus-valued entity: [-PLURAL, -PARTICIPANT] (where PARTICIPANT corresponds to 1st and 2nd person). A version of this theory holds that 3SG is absence of number and person; see Harley and Ritter (2002), Nevins (2007) for discussion. The plural of the 3PL feature in Hebrew can be explained by the semantic plurality of the reference: people in general including me and you. But given the inclusion of speaker and addressee, the 3rd person feature still has no semantic motivation.

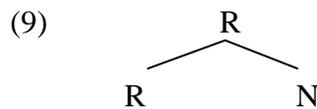
Furthermore, we propose that in Finnish, Brazilian Portuguese, and other languages with a null generic subject pronoun and agreement, i.e. the partial pro-drop languages, the generic pronoun has to have some person and number feature value because the agreement features of T have to be assigned a value. The favoured feature values are 3SG because these

⁴ Thanks to Peter Arkadiev for discussion of the Russian case with us. It seems that the 3PL form is not used in the pure inclusive sense in Russian. Instead, they tend to rely on the 2^{SG} pronoun.

are the minimal feature values.⁵ That is to say, the features are due to a morphological requirement.

In languages that have no agreement, morphology there is no such requirement. This means that the generic pronoun in, for example, Thai has no ϕ -features. A version of this idea is articulated in Phimsawat (2011). A corollary is that the pronoun must be null: Because it has no ϕ -features, there is nothing to spell out. The pronoun has minimal specification, hence minimal form (null), hence maximally inclusive reference.

Phimsawat (2011) proposes that the structure of the generic pronoun is (9):



Here, R is a referential feature, a property of all arguments (an alternative label would be D, in more or less the sense of Longobardi (1994)). In the case of referential arguments, the value of R is a referential index, regarded as a syntactic feature. R can be assigned a value freely, or by virtue of anaphoric binding, or by operator-binding. In generic arguments, R is bound by a generic operator in C, an adverbial operator GENERALLY_x .

It is not quite true, though, that the inclusive generic pronouns discussed so far, even the ones in Thai, have no restricting features: In all the examples given their reference is restricted to humans. Is this a defining characteristic of the generic inclusive pronoun? Since the reference of the pronoun always, by definition, includes the speaker, it has to include humans in its reference. Is it the case, though, that it cannot also include non-humans?

We may note first of all that it can include other conscious beings than humans, such as technologically advanced aliens from outer space or fictional talking animals. The reference of *one* in (9) could include such beings. When we say ‘human’ in the following, we actually refer to conscious beings more generally.

To test whether generic inclusive pronouns can include reference to non-humans, we need to select a predicate which can be applied to humans and non-humans. All the examples so far have had predicates with a human bias: ‘be in love’, ‘finish a BA’, ‘eat vegetarian food in October’, etc. A predicate which can apply to humans, animals and plants is ‘grow’. We have tested a number of languages using this predicate. The question is whether a sentence such as the following can refer to humans only, or if it can refer to humans *and* plants.

- (10) *thâa* \emptyset *dâayrâb* *khwaamrâk* *khwaam?awcaysây* \emptyset *kôo* *cá* *too* *rew*. [Thai]
 if get love care then FUT grow fast
 ‘If ones (animals, plants included) get love and care, ones will grow up faster.’

The translation into English clearly can only refer to humans. The generic pronoun *one* can only include humans in its reference. But the Thai sentence can refer to humans and plants.

It turns out that there is some interesting cross-linguistic variation. According to the data we have at this point, the following languages are like Thai in that the null inclusive generic pronoun can include plants along with humans and animals in its reference: Chinese (Mandarin and Taiwanese), Korean, Japanese, Sinhala, Vietnamese.⁶ We give an example

⁵ As mentioned, 2^{SG} is another option in many languages, which we ignore in this paper. Incidentally, Thai is a language where the 2^{SG} pronoun cannot be generic (see Gruber 2013) for a survey of the 2SG generic pronoun across languages).

⁶ Thanks especially to Seiki Ayano, Shin-Sook Kim, and Ji Young Shim for discussion of the Japanese and Korean facts, which made it clear to us what the general pattern is, among the radical pro-drop languages.

from Mandarin Chinese and one from Korean. Both mean ‘If one gets a lot of nutrition, one will grow fast’, but crucially, they can include humans as well as animals and plants in their reference.⁷

(11) a. Rúguǒ n ǐng hu ò d é g èng duō de y íng y ǎng, n à me hu ì zh ǎng de g èng ku ǎ.
 if can get of more of nutrition, (that) (will) grow of more fast
 [Mandarin Chinese]

b. yeongyangpwun-ul seopchwiha-myeon, ppali calaŋ-ta [Korean]
 nutrition -ACC take -if quickly grow.PRES DECL

But in the following languages the null generic 3rd person pronoun can only include humans: Brazilian Portuguese, Finnish, Hebrew, Icelandic, Polish, Tamil.

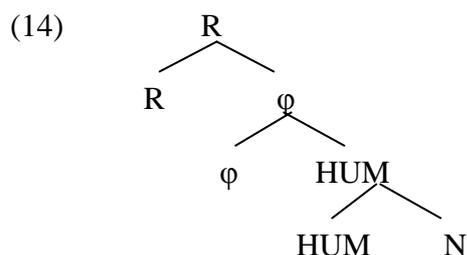
(12) im meqablim harbe ahava ve maym az gdelim maher.⁸ [Hebrew]
 if receive.3PL much love and water then grow.3PL faster
 ‘If one gets much love and water, one will grow faster.’

(13) Sit ä kasvaa nopeammin jos saa paljon ravintoa. [Finnish]
 EXPL grow.3SG faster if get.3SG much nutrition
 ‘One grows faster if one gets plenty of nutrition’

We can therefore maintain that the structure and composition of the null generic pronoun in Thai is as in (9) above. And we propose, as a tentative hypothesis, that this is also the case in the other languages where the pronoun is not restricted to humans.

The data we have at this point suggest that a crucial difference between the languages where the inclusive pronoun is all-inclusive and the ones where it is restricted to humans is subject-verb agreement: The former set of languages lack subject-verb agreement, while the latter set of languages all have it. We have no clear idea, at this point, how to explain this correlation, so we put it aside, for future research.

However, we propose that the featural make-up of the null generic pronoun in the languages where it can only refer to humans is (14).⁹



⁷ Sigurðsson and Egerland (2009: note 13) mention that the null arbitrary (not generic) impersonal subject can refer to animals provided that the predicate is animal-specific. They provide the following example:

(i) Þá var hneggja ð áhesth úsinu. [Icelandic]
 then was neighed in the.stable
 ‘Some X then neighed in the stable.’

⁸ Thanks to Ur Shlonsky for the example.

⁹ The relation between the human feature and the (other) ϕ -features is an interesting issue. We assume it is lower than the other ϕ -features, since it restricts the values that the other features can have, particularly gender and person. For example, 1st person requires [HUM].

The analysis here and in (9) presupposes that there is a categorial feature N. This is by no means obviously true. An alternative is that what makes a head or phrase nominal is that it encodes functional features such as number, gender/class, animacy, humanness, person, and honorific status. However, the fact that there is a type of argument which appears not to encode any of these features, yet has the distribution and function of a nominal argument, as does the inclusive generic pronoun in Thai, is an argument in favour of the categorial feature N.

So we conclude: The null generic pronoun in Thai is featureless, apart from a nominal feature and the feature R, which is unvalued when the pronoun is merged as subject of the sentence, but which is assigned generic reference when bound by a generic operator in C. (15) is the structure of the second clause of (4).¹⁰

(15) [_{CP} GEN [_{CP} thâa [_{TP} [_{GEN} GEN, N] [mây còb trii]]]]
 if one NEG finish BA

Because the pronoun has no ϕ -features, it has unrestricted reference, including the speaker, the addressee, and everybody and everything else. To be more precise, the reference is not restricted by the features of the pronoun itself. It can still be restricted by the semantics of the predicate and other factors outside of the pronoun.

Also, because it has no ϕ -features, it is null: there are no ϕ -features to spell-out. The categorial feature N has no spelled out form, nor does the R-feature when assigned generic value.

4. Referential null arguments

The inclusive generic pronoun is one type of null argument in Thai. There are also referential null arguments, as in (16).

(16) Nít bòok wâa Ø hěn Nóy.
 Nit say that see Noy
 ‘Nit said that she saw Noy.’

The null argument in the embedded clause in (16) can be null, as it has a local enough antecedent; see Phimsawat (2011). The null argument is co-indexed with the antecedent, and as such, they are in a control relation. The antecedent can also be found in a preceding, independent sentence, specifically if it is a topic of a preceding sentence, as seen in (17), where the null subject of the second sentence is interpreted as coreferential with the subject of the first sentence; ‘|’ indicates a sentence boundary.

(17) lûuksăaw₁ (khǒŋ) ceen₂ kèŋ khamnuan | khruu₃ bòok wâa (*khăw₁)
 daughter of-GEN Jane good at calculation teacher say COMP she
 sòob lêek dâŋ khánæen súuŋsùd.
 exam maths get mark highest
 ‘Jane’s daughter is good at calculations. The teacher said that she got top marks for maths.’

¹⁰ GEN does not just bind the variable D of the argument, but also binds an event-variable in the vP: Not just the subject, but the event or situation itself is generic, not episodic (see Chierchia 1995). This is not shown in the (15)

In the absence of context, and when the inclusive generic interpretation is not an option, a null argument will be interpreted as deictic, referring to the dominant discourse participant, i.e. the speaker. In other words, without a linguistic antecedent, the speaker is the default antecedent. In other cases, when the speaker is impossible for pragmatic or grammatical reasons, the addressee is readily available as the antecedent, as in (18):

- (18) Ø tōŋ chûaj chǎn ná.
 must help me PRT
 ‘You must help me.’

Alternatively, both speaker and addressee may be available as the antecedent of the null argument, depending on the context. (19) illustrates a case where the context does allow this interpretation:

- (19) Ø maa tham ʔaahǎn kin kan mǎy.
 let cook food eat together Q PRT
 ‘Shall we cook and eat together?’

The generalisation is, an argument can be null if it has a local enough antecedent or it refers uniquely to the speaker and/or the addressee, or, as we saw in the previous section, if it has no ϕ -features (the case of the inclusive generic pronoun).¹¹ We can be a bit more specific. Following Frascarelli (2007), we assume that a null argument A is licensed, i.e. assigned a referential index, by a null Topic in the C-domain of the minimal finite sentence containing A (see Holmberg 2010a). The null Topic itself is licensed via a Topic chain linking it to a spelled out Topic argument in the discourse context. This is the definition of ‘a local enough antecedent’. In the absence of a licensed null Topic in the C-domain, the Speaker feature, always present in the highest layer of the C-domain of a finite sentence, can step in as antecedent of A, or, if this does not yield a sensible interpretation, the Addressee feature, likewise always present in the C-domain of a finite sentence, can step in and assign a referential index to A. This is adopting Sigurðsson’s (2004, 2015) theory of ‘speech features’ in the C-domain, as syntactic representations of the speaker and the addressee (without, however, accepting the theory wholesale).

5. The quasi-inclusive generic pronoun

Consider now the quasi-inclusive pronoun.

- (20) **Raw** kin cee nay đuan tulaakhom.
 we have vegetarian food in month October
 ‘We have vegetarian food in October.’

¹¹ According to Phimsawat (2011), all null arguments in Thai are ϕ -featureless, consisting of just [R, N]. Null referential arguments would have their R valued by copying the referential index of an accessible antecedent argument. A problem for this theory is the possibility of sloppy identity. In Thai even subjects allow sloppy identity, as in (i).

(i) cɔn bɔk wâa lûuksâaw khǒŋ khǎw kɛŋ khàntasàat lé ceen bɔk wâa ___ kɛŋ phaasâa ʔaŋkrít
 John say that daughter of him good maths and Jane say that good language English
 ‘John said that his daughter is good at maths, and Jane said that *John’s daughter/ Jane’s daughter/ Jane* is good at English.’

The sloppy identity reading (‘Jane’s daughter’) cannot be the result of copying a referential index as there is no such referential index in the sentence. Sloppy identity is compatible with an NP deletion analysis of pro-drop (Tomioka 2003). We leave this issue open in this paper.

The pronoun here can have referential interpretation, referring to the speaker and some person or group associated with him/her, which may or may not include the addressee. In the right context, typically if the preceding sentence has a Topic which refers to ‘us’ (the speaker and his/her associates), the pronoun can then be null, licensed by a null Topic in the C-domain, which itself is linked to the Topic in the preceding sentence. But it can also have a generic reading, as when uttered by a Thai person to a foreigner (who understands Thai), meaning ‘Thai people in general have vegetarian food in October’. It is quasi-inclusive, not all-inclusive, as it does not (necessarily) include the addressee. Its reference is therefore restricted; it has ϕ -features: [+1, -2], if we define person in terms of two binary features [± 1 , ± 2], a PL feature, and in Thai also an honorific status feature.¹² Assuming that it has the structure [R [ϕ N]], the R-feature can be assigned generic value by the GEN operator, as we postulated in the case of the inclusive generic pronoun. However, we will return to this issue below, in section 7.

A criterial difference between the specific and the generic reading is that the generic reading allows exceptions (see Moltmann 2006). Under the specific reading, say, if *raw* in (20) refers to a female speaker and her husband, it would be false if one of them would actually not eat vegetarian food in October. But under the generic reading, (20) would be true even if some Thai people don’t eat vegetarian food in October (in fact it could be true even if most Thai people, including the speaker, don’t eat vegetarian food in October).

The quasi-inclusive pronoun can be null in the right context. If there are two occurrences of the quasi-inclusive generic pronoun in the same sentence uttered out of the blue, one in the matrix sentence, one in an embedded sentence, then the generic pronoun in the matrix subject position must be overt, the one in the embedded sentence must be null.

- (21) *(*raw*) kin cee nay đuan tulaakhom lăŋ Ø thamboonsăjbàat.
 we have veg.food in month October after offer food to monk
 ‘We have vegetarian food in October after offering food to monks.’

This is a case of an extended notion of control; see Phimsawat (2011). It falls under the principle that a null argument needs a local enough antecedent, although the mechanism is arguably not the same as in the case of licensing across independent sentences in discourse.

Furthermore, if the quasi-inclusive pronoun has an antecedent in a preceding sentence, it can be null.

- (22) raw mii p̀rapeniipàtibàt m̀aakmaay th̀i muanthay | (**raw**) kin cee
 we have tradition lots in Thailand we have vegetarian food
 nay đuan tulaakhom lăŋ Ø thamboonsăjbàat.
 in month October after offer food to monks
 ‘We have lots of traditions in Thailand. We have vegetarian food in October after offering food to monks.’

That is to say, the quasi-inclusive generic pronoun behaves much the same as its referential counterpart: It can be (sometimes must be) null when controlled. A special case of this is when it is controlled by a null topic.

At this point we will turn to the exclusive generic pronoun, and then come back to the quasi-inclusive one.

¹² Pronouns in Thai are not generally marked for number, but they are marked for honorific status; see Iwasaki and Ingkaphirom (2005: 49-57).

6. The exclusive generic pronoun

The generic exclusive reading excludes the speaker and the addressee. As mentioned, a characteristic property of the exclusive generic sentence in Thai is that the subject is optionally null, while in Italian, it is obligatorily null.

(23) a. Thîi mûubâan nîi ((phûak-)khăw) mây kin núa kanleey.¹³
At village DEM they NEG eat meat at all
'In this village they don't eat meat at all.'

b. (*Loro) parlano molte lingue in India.
they speak many languages in India

Another characteristic property of the exclusive generic pronoun is that it needs a locative adverbial. Remove the locative adverbial from (23a,b), and the subject can only be interpreted as referential. Furthermore, if uttered out of the blue, the pronoun must then be overt in Thai, and can be overt in Italian (if it is a shifted Topic or is contrastive).

(24) a. Phûak-khăw mây kin núa kanleey.
They NEG eat meat at all
'They don't eat meat at all.'

b. (Loro) parlano molte lingue.
They speak many languages

This is not something unique to pro-drop languages. Basically the same situation holds in English, and probably many other languages, as discussed in Cabredo Hofherr (2003) and especially Brody (2013). Brody notes that, for example, (25a) cannot have the exclusive generic (or as he calls it 'impersonal/universal') reading, while (25b) can.

(25) a. They like to take a nap in the afternoon.
b. In Italy, they like to take a nap in the afternoon.

The adverbial can also be a temporal adverbial of the sort which denotes a temporal stage, such as 'in the Middle Ages'.

(26) In the Middle Ages they generally died young.

Compare also (27):

(27) (In Italy) people like to take a nap in the afternoon.

The noun *people* can have a generic reading with or without a locative adverbial. Comparison of (25a,b) and (27) offers a possible clue to why the 3PL pronoun in (23), (25b), and (27) can be interpreted as generic: the locative (or temporal) adverbial licenses a silent argument 'people'; see Brody (2013). In the following, we will exploit this idea both for the exclusive and the quasi-inclusive case.

¹³ *Khăw* is an abbreviated version of *phûak-khăw* commonly used in spoken Thai. *phûak-* encodes plurality. A context is essential to determine whether *khăw* is 3SG or 3PL.

We may note first that this will explain why the pronoun is optionally null in (23a). There is, or can be, a covert noun ‘people’ in the construction which binds the pronoun, which is then null (unless it is emphasized/contrastive). In the absence of the covert noun, the pronoun will be overt. We will assume that ‘silent people’ (Brody’s term) is a null Topic, preceded by a scene-setting locative or temporal PP (of the right kind).

- (28) [CP [thîi mùubâan níi] [TOPP khon_i Top [TP Ø_i mây kin núa kanleey]]]
 at village DEM people they NEG eat meat at.all

(23a) would then be derived by a rule of ‘people deletion’: the Topic is not phonologically spelled out. Normally, as discussed in section 2, following Frascarelli (2007), referential null arguments are bound/controlled by a null Topic, which itself is interpreted by being linked to a discourse Topic. In the case of (23a/28), the null Topic ‘people’ is licensed by the locative PP, by a grammatical mechanism which we do not entirely understand.¹⁴

Brody (2013) remains entirely vague as regards the formal implementation of the idea that exclusive generic expressions involve ‘silent people’. He concludes that ‘silent people’ is “present only semantically”, with no further attempt to formalise it; a disappointingly vague hypothesis. He rejects the idea ‘silent people’ would head a DP modified by the locative PP, so that, for example, (25b) would be derived from (29) by deletion of *people*.

- (29) [DP people in Italy], they like to take nap in the afternoon

An argument against this analysis is that the putative rule of *people* deletion cannot apply to such DPs when they are in subject or object position. We cannot, for example, derive (30b) from (30a).

- (30) a. People in Italy like to take a nap in the afternoon.
 b. *In Italy like to take a nap in the afternoon.

But this argument does not affect our analysis (28), where ‘people’ is a null Topic licensed by a scene-setting adverbial. An argument against this analysis, though, as the only possible configuration licensing ‘silent *people*’ is that the locative (or temporal) expression need not be in sentence-initial position the way it is in (23a) or (25b). In the Italian example (23b), for example, the locative PP is quite clearly an adjunct to VP.

Consider (31), though, the Thai counterpart of (23b).

- (31) khăw phûut kan lăay phaasăa thî ?india. [Thai]
 they speak together many language at India
 ‘They speak many languages in India.’

Here the pronoun, which can be generic, has to be overt. This indicates that the analysis in (28) might be right after all: In (31), because the locative PP is not in the C-domain, a null Topic ‘people’ is not licensed, and for this reason, the pronoun has to be overt. So how is the generic reading then effected in (31)? Another interesting property of (31) is that the adverb *kan* ‘together’ is required for the generic reading. Without it, the pronoun will

¹⁴ Gruber (2013) articulates a theory of indexical pronouns, including the 2SG generic pronoun, where they all require spatial anchoring, by an overt or covert spatial expression. This theory could conceivably be extended to the exclusive generic case, even though it is not indexical (i.e. 1st or 2nd person).

be interpreted as referential. We do not profess to understand how this mechanism works, so we leave it for future research.

Brody (2013) concludes, from his discussion of exclusive generic *they* in English that this is not a different, generic *they*, but the ordinary referential pronoun *they*, which gets its generic reading because it is bound by the silent generic expression *people*. We are happy to extend this conclusion to Thai as well, and probably much more generally.

Consider Italian; we repeat the example (23b) as (32):

- (32) Parlano molte lingue in India.
 speak.3PL many languages in India
 ‘They speak many languages in India.’

Recall that Italian, by hypothesis, has a definite feature in T with the effect that null 3SG pronouns can only be interpreted as definite/referential. This explained why Italian needs an overt inclusive generic pronoun *si*. Then how come (32) can have a null 3PL pronoun interpreted as (exclusive) generic? We suggest this is because the null 3PL subject in (32) is not a null version of the definite pronoun *loro* ‘they’, but ‘silent people’, with deletion licensed by the locative adverbial. Again, we will leave the formal details of this analysis for future research.

7. The quasi-inclusive pronoun revisited

If the generic interpretation of the exclusive generic pronoun is due to a silent noun ‘people’, we might wonder whether this silent noun is also part of the syntax of the other generic pronouns.

We can safely discard the idea that it would be involved in the case of the inclusive generic pronoun. To begin with, as we demonstrated above in section 3, in Thai and a number of other radical pro-drop languages the reference of the inclusive generic pronoun can include not just people but also, for example, plants. Second, in most of the partial pro-drop languages, where the null inclusive pronoun does refer exclusively to people, the agreement on the verb is singular, not plural as we would expect if the subject is ‘people’. The only clear exception that we are aware of is Hebrew. In this language we may indeed consider the possibility that there is a silent noun ‘people’ involved.

For the quasi-inclusive generic pronoun, though, we may consider the possibility that it is involved. More specifically, we may consider the possibility that the quasi-inclusive generic pronoun is an adnominal construction ‘we people’, with ‘people’ deleted, under essentially the same condition as in the case of the exclusive generic pronoun. It is certainly typical of the quasi-inclusive generic construction, too, to have a locative restrictor (or temporal, of the right kind). For example, corresponding to (33a), there would be (33b), where the PF of (33b) is derived by deletion of the noun ‘people’.

- (33) a. raw khon thai mii p̀rapeeniipàtibàt màakmaay.
 we Thai have tradition lots
- b. [raw ~~khon~~] mii p̀rapeniipàtibàt màakmaay thii muanṭhay.¹⁵
 we people have tradition lots in Thailand
 ‘We have lots of traditions in Thailand.’

¹⁵ To construct a well-formed sentence, the deletion of *khon* ‘people’ is obligatory.

The deletion of *khon* ‘people’ would be licensed by the locative adverbial, as in the other cases discussed above (leaving the details for future research).

This analysis would be particularly natural in a theory of pronoun structure where definite pronouns always have the structure [D NP], where D is spelled out as the pronominal form and the NP component is normally null, but can be spelled out as in *we linguists*, *you students*, etc.; Panagiotidis (2002), Elbourne (2008). The quasi-inclusive generic pronoun would be *we people*, but with the NP component deleted/not spelled out. This analysis cannot easily be extended to the exclusive pronoun because *they people*, *they students*, etc. is an ungrammatical construction in most, though not all, languages (Georg Höhn, p.c.). Furthermore, the variation between (28) and (31) would not easily be captured under such an analysis.

The analysis of the exclusive and quasi-inclusive generic pronouns as derived by ‘people deletion’ raises the issue what role, if any, is played by the generic operator GEN, which we postulated, following much work particularly on the semantics of generic expressions (Carlson and Pelletier 1995, Moltmann 2006). In the case where the postulated, deleted noun ‘people’ is a component of a DP argument, we may assume that it is bound by a generic operator in CP. But this would presumably not be the case where ‘people’ is itself a null topic, binding a null argument in TP. In that case, it would seem that the generic reading is an effect of the null Topic itself. We leave this issue as well for future research.

8. Conclusions

Our conclusions as regards the features of the three generic pronouns are:

- (1) The inclusive generic pronoun in Thai (and in radical pro-drop languages more generally) has no ϕ -features. It is made up of nothing but an abstract noun and an R (or D) head assigned generic value by the generic operator GEN in the C-domain. This accounts for its interpretation (maximally general) and its form (null).
- (2) The quasi-inclusive generic pronoun is derived from an adnominal construction ‘we people’ by deletion of ‘people’.
- (3) The exclusive generic pronoun is also derived by ‘people deletion’. One construction where this is commonly found is where ‘people’ is a null Topic in the C-domain binding a pronoun in TP. In Thai, the pronoun is null in this case.

References

- Barbosa, P (to appear). *Pro* as a minimal NP: towards a unified approach to pro-drop. To appear in *Linguistic Inquiry*.
- Biberauer, T., Holmberg, A., Roberts, I., & Sheehan, M. (2010). *Parametric syntax: Null subjects in minimalist theory*. Cambridge: Cambridge University Press.
- Blevins, J. (2003). Passives and impersonals. *Journal of Linguistics* 39: 473-520.
- Brody, M. (2013). Silent People: The pseudo-impersonal. In Johan Brandtler, Val éria Moln á, Christer Platzack (eds.), *Approaches to Hungarian: Volume 13: Papers from the 2011 Lund conference*. John Benjamins Publishing Company.
- Cabredo Hofherr, P. (2003). ‘Arbitrary readings of 3pl pronominals’. In M. Weisgerber (ed.) *Proceedings of the Conference “sub7— Sinn und Bedeutung”, 7th Annual Meeting of the Gesellschaft für Semantik*. Arbeitspapier Nr. 114, FB Sprachwissenschaft, Konstanz University.
- Carlson, G. & Pelletier, P. (eds.) Carlson. (1995). *The Generic Book*. Chicago University Press.

- Chierchia, G. (1995). Individual level predicates as inherent generics. In Carlson & Pelletier (eds.) *The Generic Book*. Chicago University Press, 176-223.
- Cinque, G. (1988). On *si* constructions and the theory of *ARB*. *Linguistic Inquiry* 19: 521-81.
- Dobrovie-Sorin, C. (1998). Impersonal *se* constructions in Romance and the passivization of unergatives. *Linguistic Inquiry* 29, pp. 399-437.
- D'Alessandro, R. (2007). *Impersonal si constructions*. Berlin/New York: Mouton de Gruyter.
- Elbourne, P. (2008). The interpretation of pronouns. *Language and Linguistics Compass*: 119–150.
- Frascarelli, M. (2007). 'Subjects, topics and the interpretation of referential pro: An interface approach to the linking of (null) pronouns'. *Natural Language and Linguistic Theory* 25: 691–734.
- Gruber, B. (2013). *The spatiotemporal dimensions of person. A morphosyntactic account of indexical pronouns*. PhD dissertation. LOT, Utrecht University.
- Harley, H. & Ritter, E. (2002). Person and number in pronouns: a feature-geometric analysis. *Language* 78:482-526.
- Holmberg, A. (2005). 'Is there a little pro? Evidence from Finnish'. *Linguistic Inquiry* 36: 533–564.
- Holmberg, A. (2010a). 'Null subject parameter'. In T. Biberauer et al. *Parametric Variation: Null Subjects in Minimalist Theory*. Cambridge, UK: Cambridge University Press. 88–124.
- Holmberg, A. (2010b). 'The null generic subject pronoun in Finnish: a case of incorporation in T'. In T. Biberauer et al. *Parametric Variation: Null Subjects in Minimalist Theory*. Cambridge, UK: Cambridge University Press. 200–230.
- Holmberg, A., Nayudu, A., & Sheehan, M. (2009). 'Three partial null-subject languages: a comparison of Brazilian Portuguese, Finnish, and Marathi'. *Studia Linguistica* 63:59–97.
- Holmberg, A. & Sheehan, M. (2010). 'Control into finite clauses in partial null-subject languages'. In T. Biberauer et al. *Parametric Variation: Null Subjects in Minimalist Theory*. Cambridge, UK: Cambridge University Press. 125–152.
- Iwasaki, S. & Ingkaphirom, P. (2009). *A Reference Grammar of Thai*. Cambridge: Cambridge University Press.
- Krzek, M. (2013a) Interpretation and voice in Polish SIE_ę and –NO/–TO constructions. In I. K. Chahine (ed.), *Contemporary Studies in Slavic Linguistics*. Amsterdam: John Benjamins 185-198.
- Krzek, M. (2013b). Generic subjects and Voice in Polish impersonal constructions. In C. S. Rhys, P. Iosad, and A. Henry (eds.), *Microvariation, Minority Languages, Minimalism and Meaning: Proceedings of the Irish Network in Formal Linguistics*'. Cambridge: Cambridge Scholars Publishing. 186-206.
- Longobardi, G. (1994). Reference and proper names. *Linguistic Inquiry* 25: 609-665.
- Manninen, S. & Nelson, D. (2004). What is a passive? The case of Finnish. *Studia Linguistica* 58:212-251.
- Moltmann, F. (2006). Generic one, arbitrary PRO, and the first person. *Natural Language Semantics* 14: 257–281.
- Nevins, A. (2007). The representation of third person and its consequences for person-case effects. *Natural Language & Linguistic Theory* 25 (2), 273-313.
- Panagiotidis, E. P. (2002). *Pronouns, clitics and empty nouns*. Amsterdam: John Benjamins.
- Phimsawat, O. (2011). *The syntax of pro-drop in Thai*. Ph.D. dissertation, Newcastle University.
- Sigurðsson, H. A. & Egerland, V. (2009). Impersonal null subjects in Icelandic and elsewhere. *Studia Linguistica* 63: 158–185.

- Sigurðsson, H. A. (2004). The syntax of Person, Tense, and speech features. *Italian Journal of Linguistics* 16:219–251.
- Sigurðsson, H. A. (2009). Impersonal null-subjects in Icelandic and elsewhere. *Studia Linguistica* 63: 158-185.
- Sigurðsson, H. A. (2015). *About pronouns*. Ms. Retrieved from <http://ling.auf.net/lingbuzz/001593> on 11/4/15.
- Tomioka, S. (2003). The semantics of Japanese null pronouns and its cross-linguistic implications. In K. Schwabe and S. Winkler (eds.). *The interfaces: deriving and interpreting omitted structures*. Amsterdam: Benjamins, 321-40.

*Anders Holmberg
School of English Literature, Language and Linguistics
Percy Building, Newcastle University, Newcastle upon Tyne
NE1 7RU,
United Kingdom.*

anders.holmberg@ncl.ac.uk

*On-Usa Phimsawat
Department of Western Languages
Faculty of Humanities and Social Sciences
Burapha University
169 Long-Hard Bangsaen Road,
Saen Sook Sub-district,
Mueang District, Chonburi 20131,
Thailand*

onusa@buu.ac.th

*AFFECT IN CHINESE AND KOREAN SPOKEN NARRATIVES*¹

VIVIAN LEE
(SOAS, University of London)

LU LU
(SOAS, University of London)

Abstract

This paper details a pilot study which was conducted to investigate the Affect sub-category of the appraisal system in Chinese and Korean speakers' L1 and L2 narrations. Affect is concerned with the expression of interpersonal meaning and has been widely applied to discourse analysis in various genres. It looks at the registering of positive and negative feelings: do we feel happy or sad, interested or bored? Emotive vocabulary in Chinese and Korean L1 (first language) and L2 (second language) narrations has been relatively less explored, and the researchers thus attempt to study the affect system based on such speakers' narrations. For the pilot study, narratives were elicited using a short film. The researchers manually annotated the transcribed data for expressions belonging to the discourse meaning of dis/inclination, un/happiness, in/security, and dis/satisfaction. Preliminary results show that Chinese and Korean speakers' L1 data do not show a strong influence on the L2 data in general, and that fewer emotive expressions are likely to be found in the L1 data than the L2 data. The results of this exploratory study can lay a foundation for the future main study, which will compare Chinese and Korean speakers' L1 and L2 narrations, in addition to look into any possible linguistic and cultural influences.

1. Introduction

While emotions are embodied by all human beings, differences have been found in the way they are expressed by speakers with different L1s. Various approaches to emotion analysis have been proposed, such as in studies by Wierzbicka (1999) and Pavlenko (2002, 2008a, 2008c). Affect, a sub-category of the Appraisal System (Martin and White 2005), is a useful theoretical model to study the emotive dimension of meaning. While studies have been carried out on European languages (Pavlenko 2008c; Segalowitz et al. 2008), to date limited research has compared L2 English learners' emotive expressions in spoken corpora between two East Asian languages: Chinese and Korean. This exploratory article thus attempts to open up a promising research venue and encourage research towards the investigation of emotive expressions from a corpus-based approach.

This article is structured as follows: Section 2 gives background to the study, including an outline of Martin and White's (2005) proposal on the sub-category of Affect, and reviews its applications in second language acquisition and narrative analysis, with special references to the Korean and Chinese languages. Section 3 details the methodology used for the pilot study. Section 4 reveals preliminary results, and Section 5 concludes the article.

¹ The authors are grateful to their study participants for their valuable time and help, and the audience of AE-LINK 2 for their useful comments. We would also like to express gratitude for the support of the MOE Key Project of Key Research Institute of Humanities and Social Sciences at Universities of China (ref. 14JJD740014).

2. Background

2.1. The Appraisal System and the sub-category Affect

Martin and White's (2005) Appraisal system is an analytical model that attempts to investigate how attitudes and dialogic engagement are explicitly and/or implicitly encoded in texts and the gradability of attitudinal meanings and engagement values. Among these research areas, we refer to the sub-system that examines how feelings are construed in texts as Attitude. The discourse semantic dimension that encodes emotions is known as Affect, the focus of the paper.

The sub-category of Affect concerns the registering of “positive and negative feelings: do we feel happy or sad, confident or anxious, interested or bored”? (Martin and White 2005: 42). In Martin and White's (2005) monograph, they further categorised the category of Affect into four different types of realisations: dis/inclination, un/happiness, in/security, and dis/satisfaction.

Among the four kinds of Affect, un/happiness seems to be the first sub-category that one can think of. It refers to “the moods of feeling happy or sad, and the possibility of directing these feelings” at “the phenomenon responsible for that emotion” by “liking or disliking it” (Martin and White 2005: 46-49). For instance, *the captain felt sad/happy* (Martin and White 2005: 49). The dis/inclination set of discourse semantics is the feelings of yearning for or fear. Examples could be *kewang* ‘yearn for’ in Chinese and *twulyepta* ‘fear’ or ‘dread’ in Korean. The in/security category involves the feeling of “peace and anxiety in relation to our environs” (Martin and White 2005: 49), such as the lexical item *startled* and their Korean and Chinese correspondences *jingxia* and *nolata*, ‘startled’ or ‘surprised’. The last sub-category dis/satisfaction covers human being's feeling of “achievement and frustration in relation to the activities in which we are engaged, including our roles as both participants and spectators” (Martin and White 2005: 50). Such instances as *satisfied* and *fed up with* in English, *manyi* ‘satisfied’, *youxingqu* ‘feel interested’ in Chinese, *sinkihata* ‘marvel at’, *hansimhata* ‘pathetic’ in Korean best illustrate this sub-category. Table 1 lists the working definitions of the four sub-categories of Affect, and their Korean and Chinese examples collected for this project.

Table 1. An overview of the sub-categories of Affect in this study (based on Martin and White 2005)

Categories	Sub-categories & codes	Working definitions	English examples	Korean examples	Chinese examples
dis/inclination	inclination [INCL]	the feelings of being inclined to do something	long for; miss	<i>hungmilul kacta</i> ‘to have interest (in/to do something)’	<i>manxiang</i> ‘be quite willing to’; <i>yueyueyushi</i> ‘be eager to do something’

	disinclination [DINCL]	the feelings of unwillingness or fear	wary; fearful	<i>twulyepta</i> 'fear' or 'dread' <i>mwusepta</i> 'fear/afraid'	<i>mianqiang</i> 'be reluctant'; <i>konghuang</i> 'be terrified'
un/happiness	happiness [HPY]	the moods of feeling happy	cheerful; love	<i>culkepta</i> 'enjoyable/fu n' <i>kippehata</i> 'be happy'	<i>gaoxing</i> 'be happy'; <i>xiyue</i> 'be pleased'
	unhappiness [UHPY]	the moods of feeling sad	sad; in tears	<i>koylepta</i> 'be in (emotional) pain/painful'	<i>kuaiku-la</i> 'nearly in tears'; <i>tongku</i> 'sad'
in/security	security [SECU]	the feelings of peace	together; confident	<i>casinmanman</i> <i>hata</i> 'be full of confidence'	<i>xinxin</i> <i>shizu</i> ; 'full of confidence'
	insecurity [ISECU]	the feelings of anxiety	uneasy; surprised	<i>nolata</i> 'startled' or 'surprised' <i>tanghwangha</i> <i>ta</i> 'be flustered'	<i>hen jingya</i> 'astonished' ; <i>hao kepa</i> 'dreadful'
dis/satisfaction	satisfaction [STS]	the feelings of achievement	pleased; thrilled	<i>sinkihata</i> 'marvel (at)'	<i>you xingqu</i> 'have interest in'; <i>hen jingcai</i> 'amazing'
	dissatisfaction [DSTS]	the feelings of frustration	sick of; fed up with	<i>hansimhata</i> 'pathetic' <i>mwusihata</i> 'look down on'	<i>mei duoda</i> <i>xingqu</i> 'not have much interest in'; <i>bishi</i> 'look down upon'

As the listing shows, the realisations of Affect are various; it can be realised either as quality, such as example (1) in which happiness is used to describe the attribute of the participant, or as process, for instance the affective behavioural in (2). Likewise, for Korean, (3) is an example of disinclination as an attribute being used to describe the subject himself (Mr. Bean in this study), while (4) describes the young boys' dissatisfaction with Mr. Bean. Furthermore, the lexical elaborations in the above tables are not exhaustive and this paper attempts to investigate the Affect that is typical of the Korean and Chinese fictional narrations (see Section 4 for details). It is also hoped to finer classify the sub-categories of Affect.

- (1) Ta hen kaixin.
He very pleased
'He is very pleased.'
- (2) Xiao nanhai bishi handou xiansheng.
little boy despise Bean Mr
'The little boy despised Mr Bean.'
- (3) Wancenhi kep-ul mek-ess-supnita.
completely fear-OBJ² eat-PST-DECL
'(He) was extremely afraid.' (literally 'to eat fear')
- (4) Maywu mwusihay-ssul keskath-supnita.
very look.down-PST seems-DECL
'It seems like (they) looked down on him a lot.'

2.2. Emotions in L1 and L2

Emotion in speakers and its relation to second language acquisition has been an area of interest and is well-researched by many scholars in the field (see, for example Wierzbecka 1999; Pavlenko 2002, Pavlenko 2008a, Pavlenko 2008c, and Pavlenko 2009). Studies have found that for speakers of more than one language, words with affective meanings do not evoke the same feelings in first and second languages. For instance, Segalowitz, Trofimovich, Gatbonton and Sokolovskaya (2008) studied whether differences in L1 and L2 reflected automatic processing of the affective element of affectively valent words, and the relation between such processing and general word recognition efficiency for L2 words which lacked affective valency. It was found that the processing of affect in the L2 is less automatic. Some researchers (Kang 2003, 2006; Pavlenko 2008b; Leichtman, Wang and Pillemer 2003; Chang and McCabe 2013; Anolli et al. 2008) made further discussions to see if speakers tend to transfer their L1 emotive expressions into their L2. Pavlenko (2008a) argues for the importance for models of the bilingual lexicon to acknowledge not only linguistic and cognitive but also affective aspects of the lexicon (Pavlenko 2008a: 147). Marian and Kaushanskaya (2008) studied the cross-linguistic differences in emotionality of autobiographical memories, and found that bilinguals used more emotion words when describing their immigration experiences in the second language than the first. The authors propose that bilingual speakers' expression of emotion possibly varies across languages, and that there is a connection between the linguistic and affective systems to the bilingual cognitive architecture. Panayiotou (2004) examined whether bilinguals express different emotions in their languages, and found that respondents displayed different reactions to the same story depending on the language.

Although there have been comparatively few studies on emotion words in the South Korean context, Lee's (2011) study on Korean English learners' use of emotion words found that English learners responded with various inappropriate verb forms such as 'I feel' and 'I am', while the majority of L1 English speakers responded with subjunctive forms such as 'I would feel'. In addition, L2 English learners used mostly simple and coordination sentences. It was also found that the lexical richness, measured through type/token ratio, was higher in

² In this study, we use the following abbreviations. OBJ: object marker; DECL: declarative marker; PST: past tense.

English L1 data than in English L2 data. The proportion of emotion lemmas reflects the lexical richness or the diversity of the emotion words. Lastly, it was found that L2 English learners' responses focused on a few typical adjectives such as 'happy', 'angry' and 'scared'.

As for the narratives in Chinese, Chang and McCabe's (2013) study investigated the evaluative strategies among Mandarin Chinese-speaking children from Taiwan and English-speaking children from the United States. One of the evaluative strategies is the descriptions of internal emotional states, which belong to the Affect system in our study. Chang and McCabe (2013) reported that Taiwanese children, overall, produced more descriptions of internal emotional states than the American children, though the average percentage of evaluative comments was lower than their counterparts in America. In another study, Tsai et al. (2005) did an analysis on emotive expression among Chinese Americans (whose household languages were varieties of Chinese such as Mandarin, Taiwanese or Cantonese), compared with European Americans. In their empirical study, they found the different ways they described emotional experiences: less acculturated Chinese Americans used more somatic (e.g. *dizzy*) and social (e.g. *friend*) words to describe emotions than European Americans when they were speaking English. Inspired by Martin and White's (2005) Appraisal system, Xu (2013) conducted a comprehensive investigation regarding Chinese L2 English learners' appraisal patterns. For the Affect system, his examination showed that the L2 English learners in China were inclined to overuse Affect, compared with the English native speakers in question, and that Chinese English L2 learners, in general, tended to narrate in a more direct and less fleshed-out way than their native English respondents.

As such, it is evident that differences exist in the narration style and techniques of Chinese and Korean speakers. The current study will seek to explore preliminary emotion in narratives produced by Chinese and Korean speakers in their L1 and L2, to identify characteristics between the two languages.

Considering the existing research in this area, the following questions are posed in this study.

- (1) What affective expressions do Korean and Chinese L1 speakers use in their fictional narrations?
- (2) What affective expressions do the speakers use in their L2 fictional narrations?

3. Methodology

This section will detail the methodology used for the pilot study. We attempt to address the above questions with a corpus-based approach. This section will describe the narrative story we chose, the way we compiled the corpus and the coding scheme.

The researchers used *Mr Bean in the Swimming Pool*, a five-minute video clip, to elicit data (Pavlenko 2008c). Ten Korean students and ten Chinese students were asked to narrate the story as it played. As the purpose of the current pilot study was to elicit data for preliminary findings, five speakers of each group were asked to narrate in their L1, and the other five in their L2, English. They were then asked four follow-up questions (see Appendix) in the language they were asked to narrate in. Their responses to the interview questions were also given in the same language they had been asked to narrate in i.e. Chinese, Korean and English respectively. Having manually annotated the transcriptions from the obtained spoken data, researchers conducted an analysis based on the Affect section of the Appraisal System, with subgroups of dis/inclination, un/happiness, in/security and dis/satisfaction.

3.1. Materials

Participants were shown a short film with no dialogue but with soundtrack. The film is an extract from *The Curse of Mr Bean*, and is a part of Act 1, which is available on the internet as *Mr Bean at the Swimming Pool*. In the video, the comical character Mr Bean wants to go on the children's slide at a swimming pool. The lifeguard tells him off, and Mr Bean has no choice but to leave the children's area. He goes to the adult's section and spots a high diving board. Once he is up the steps though, Mr Bean becomes frightened of the height and changes his mind. However, there is no turning back for Mr Bean as two young boys are behind him, waiting for their turn. In an attempt not to appear cowardly, Mr Bean makes a few comical tries to dive off, but failing each time. The video ends where one of the young boys stamps on Mr Bean's hand, who is now hanging off the board with that hand. This leads Mr Bean to fall into the swimming pool. A version of the video was used previously by Pavlenko (2008c) and was chosen for its usefulness in eliciting affect-related vocabulary, as well as its easy accessibility.

3.2. Corpora

In order to compare and contrast Korean and Chinese L2 English learners' narrative strategies and the possible influences from their L1, ten Korean students and ten Chinese students were recruited on a voluntary basis. They were asked to narrate *Mr Bean at the Swimming Pool* as the video clip played. Five speakers of each group were asked to narrate in their L1, and the other five in their L2, English. They were then asked four follow-up questions (see the Appendix), based on the story they had just watched to elicit more data. Their responses to the interview questions were made in the language they had been asked to narrate in i.e. Chinese, Korean and English respectively. Although the twenty participants in this study had majored or were majoring in different disciplines, they were all from key universities in South Korea and China, and were all educated to above bachelor's degree level. Although we are aware that there could be a wide range in the difference of English proficiency among the narrators, and that proficiency level is expected to mediate the ways that L2 learners express emotion, the researchers assumed that speakers would be able to sufficiently express themselves in their L2 English. This idea also agrees with the purpose of the current study, which is designed to elicit preliminary data to identify any interesting properties for future research. As such, by employing international English tests like TOEFL or IELTS to recruit participants later on, the main project can be more comparable.

Having recorded and transcribed their narrative data, the two researchers built four comparable corpora in total. The size of each corpus is demonstrated in Table 2.

Table 2. The size of the four corpora used for the current study

Participants	L1 (characters)	L2 (words)
Korean	5,660	2,269
Chinese	4,221	2,173

3.3. Coding

For the above corpora, the researchers manually annotated the expressions in relation to the discourse meaning of dis/inclination, un/happiness, in/security, and dis/satisfaction. Martin and White's (2005) elaborations on English data aside, we coded Korean and Chinese

data based on the individual properties of the two languages while referring to the original English classifications. Table 1 in Section 2.1 elaborates their working definitions and corresponding examples in English, Korean and Chinese.

Having employed the annotation software BFSU Qualitative Coder (Xu and Jia 2011), the two researchers discussed, double-checked and cross-checked the coded data together. As for ambiguous expressions, they were classified into a separate category known as “others”, if no consensus had been reached between the two researchers. When all the checking was finished, BFSU Qualitative Explorer (Xu and Jia 2011) was used to calculate the token and type of each target category.

4. Results

The results are tabulated below. As this is an exploratory study, we are interested in identifying the affective expressions that are prominently employed in the speakers’ fictional narrations. Table 3 demonstrates the distribution of affect expression across four corpora in question: Chinese speakers’ L1 and L2 English corpora (CC and CE, for short) and Korean speakers’ L1 and L2 English corpora (henceforth KK and KE respectively).

Table 3. The distribution of Affect expressions among the four corpora

Category	Sub-category	CC	CE	KK	KE
dis/inclination	inclination	24	32	18	26
	disinclination	58	39	29	33
un/happiness	happiness	12	9	15	13
	unhappiness	2	4	1	2
in/security	security	9	10	14	5
	insecurity	23	26	28	19
dis/satisfaction	satisfaction	6	12	3	0
	dissatisfaction	10	9	23	19

We first investigated the two L2 corpora, namely the CE corpus for Chinese speakers and KE corpus for Korean narrators. The results find that the Chinese and Korean L2 learners of English demonstrate a similar tendency in expressing affects in their narrations. And such emotive categories as dis/inclination, un/happiness and dis/satisfaction in particular are more favoured by Chinese L2 English learners. However, the log likelihood ratio test, which is used to compare the fit of two groups in linguistics (Dunning 1993), found that Chinese and Korean narrators do not show significant differences ($p>0.05$) in expressing emotions when speaking English (see the statistical results in Table 4 for details).

Table 4. The log-likelihood ratio between Chinese and Korean L2 English learners

	Freq_CE	Freq_KE	LL	sig.	
dis/inclination	71	59	1.69	0.194	+
un/happiness	13	15	0.07	0.792	-
in/security	36	24	2.96	0.085	+
dis/satisfaction	21	19	0.21	0.651	+

Since Chinese and Korean are both East Asian communities, they share a lot in cultures and perceptions. We thus intend to test whether the L2 narrations show an influence from the L1. The log likelihood ratios of L1 and L2 of Chinese and Korean narrators were calculated respectively. Table 5 demonstrated that Chinese speakers highly significantly ($p < 0.01$) less prefer to express emotions when they were asked to narrate in Chinese than in English. This implies that the Chinese L1 does not influence the English L2 regarding narrative emotions, especially in un/happiness (LL=10.25; $sig=.001$), dis/inclination (LL=10.05; $sig=.002$), and in/security (LL=8.00; $sig=.005$).

Table 5. The log-likelihood ratio between Chinese speakers' L1 and L2

	Freq_CC	Freq_CE	LL	sig.	<i>p</i>	
dis/inclination	82	71	10.05	0.002	**	-
un/happiness	14	13	2.30	0.130		-
in/security	32	36	10.25	0.001	**	-
dis/satisfaction	16	21	8.00	0.005	**	-

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

The test for the influences of Korean L1 on L2 was carried out and the results were demonstrated in Table 6. Like their Chinese counterparts, the data showed that Korean speakers seem to reserve their emotions when speaking in Korean than in English. And statistical results further confirm that the affect categories of un/happiness and dis/inclination are significantly (LL=5.38; $sig=.020$) and extremely significantly (LL=33.74; $sig=.000$) favoured by Korean L2 English speakers. However, while Chinese narrators whose emotive expressions of un/happiness do not demonstrate statistical differences, Korean speakers do not show significant differences in terms of in/security and dis/satisfaction.

Table 6. The log-likelihood ratio between Korean speakers' L1 and L2

	Freq_KK	Freq_KE	LL	sig.	<i>p</i>	
dis/inclination	47	59	33.74	0.000	***	-
un/happiness	16	15	5.38	0.020	*	-
in/security	42	24	1.85	0.174		-
dis/satisfaction	26	19	3.78	0.052		-

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

In a general sense, our findings demonstrate that Korean and Chinese do not show differences in emotive expressions when speaking in their L2 of English. L1 speakers of Chinese and Korean are both inclined to use fewer emotive expressions in their native languages than in their L2 English. On the other hand, Chinese and Korean narrators show different degrees of preferences in terms of the four subcategories of affect.

5. Conclusion

This paper has detailed a pilot study which looked at Chinese and Korean speakers' use of emotive vocabulary in L1 and L2 fictional narrations. By analysing the corpus against Martin and White's (2005) Affect system, the researchers found that Chinese and Korean speakers' L1 do not show a strong influence on their L2 English narrations in general. The study found that native speakers of Chinese and Korean are both inclined to use fewer emotive expressions in their native languages than in their L2 English. Due to the exploratory nature of the this study, this paper has its limitations in that it is of a small scale; however, despite its small scale, the pilot study has made a preliminary investigation into the emotive expressions belonging to the affect system for Chinese and Korean speakers' L1 and L2. There is also a need for further analysis to be carried out on expressions of affect which did not fit clearly into Martin and White's system, and were labelled as 'other'. Further, there is a need for analysis to look at possible culture-related preferences or strategies evident in Korean and Chinese L2 English oral fictive narrations. Possible future research could include a larger-scale study which will elicit both L1 and L2 narrations from each speaker. The study could also be extended to examine other subsystems within the Appraisal System and analysis can be carried out to explore commonalities and differences among other languages.

Appendix

The four follow-up questions used following the narration in English, Chinese and Korean:

English

- (1) How does Mr. Bean feel when he first enters the swimming pool?
- (2) How does he feel when he is on top of the diving deck?
- (3) How do the young boys feel when they are watching him?
- (4) Can you describe the lifeguard's facial expression when he sees Mr. Bean hanging off the diving deck?

Chinese

- 1) 可以描述一下憨豆先生刚进游泳池的时候的心情吗?
- 2) 他站在跳台上是什么感受呢?
- 3) 当两个小男孩看到他的时候, 他们的感受是怎么样的?
- 4) 当看到憨豆先生快要从跳板上跳下去的时候, 救生员的面部表情是怎么样的?

Korean

- (1) 미스터빈이 수영장에 들어갈 때 기분이 어땠을 것 같아요?
- (2) 다이빙대에 올라가 있을 때는 미스터빈의 기분이 어땠을 것 같아요?
- (3) 남자 아이들은 미스터빈을 보고 있을 때 어떤 기분이었을까요?
- (4) 다이빙대에 매달리고 있는 미스터빈을 본 안전요원의 표정이 어땠나요?

References

Anolli, L., Wang, L., Mantovani, F. & De Toni, A. (2008). The voice of emotion in Chinese and Italian young adults. *Journal of Cross-Cultural Psychology* 39(5), 565-598.

- Chang, C. & McCabe, A. (2013). Evaluation in Mandarin Chinese children's personal narratives. In McCabe, A. and Chang, C. (eds.), 33-56.
- Dunning, T. (1993). Accurate methods for the statistics of surprise and coincidence. *Computational Linguistics* 19(1), 61-74.
- Harkins, J. & Wierzbicka, A. (2001). *Emotions in crosslinguistic perspectives*. Berlin/New York: Mouton De Gruyter.
- Kang, J. Y. (2003). On the ability to tell good stories in another language: Analysis of Korean EFL learners' oral "Frog Story" narratives. *Narrative Inquiry* 13(1), 127-149.
- Kang, J. Y. (2006). 'Producing culturally appropriate narratives in English as a foreign language: A discourse analysis of Korean EFL learners' written narratives', *Narrative Inquiry* 16(2), 379-407.
- Lee, J. (2011). Use of Emotion Words by Korean English Learners. *English Language & Literature Teaching* 17(4), 193-206.
- Leichtman, M. D., Wang, Q. & Pillemer, D. B. (2003). Cultural variations in interdependence and autobiographical memory: Lessons from Korea, China and India. In Fivush, R. and Haden, C. A. (eds.), *Autobiographical memory and the construction of a narrative self: Developmental and cultural perspectives*. Mahwah, NJ: Laurence Erlbaum Associates.
- Marian, V. & Kaushanskaya, M. (2008). Words, feelings, and bilingualism: Cross-linguistic differences in emotionality of autobiographical memories. *The Mental Lexicon* 3, 72-90.
- Martin, J. R. & White, P. R. (2005). *The language of evaluation*. Basingstoke and New York: Palgrave Macmillan.
- McCabe, A. & Chang, C. (2013). Introduction. In McCabe, A. & Chang, C. (eds.), 1-6.
- McCabe, A. and Chang, C. (eds.) (2013). *Chinese Language Narration: Culture, cognition, and emotion*. Amsterdam: Benjamins.
- McCabe, A. (1991). Editorial. *Journal of Narrative and Life History* 1(1), 1-2.
- McCabe, A. (1997). Developmental and cross-cultural aspects of children's narration, in Bamberg, M. (ed.), *Narrative development: Six approaches*. Mahawah, NJ: Erlbaum.
- Panayiotou, A. (2004). Switching codes, switching code: Bilinguals' emotional responses in English and Greek. *Journal of Multilingual and Multicultural Development* 25(2&3), 124-139.
- Pavlenko, A. (2002). Emotions and the body in Russian and English. *Pragmatics & Cognition* 10, 207-241.
- Pavlenko, A. (2005). *Emotions and multilingualism*. New York: Cambridge University Press.
- Pavlenko, A. (2008a). Emotion and emotion-laden words in the bilingual lexicon. *Bilingualism: Language and Cognition*, 11, 147-164.
- Pavlenko, A. (2008b). Narrative Analysis. In Wei, L. and Moyer, M. G. (eds.), *The Blackwell Guide to Research Methods in Bilingualism and Multilingualism*. Oxford: Blackwell Publishing Ltd.
- Pavlenko, A. (2008c). Structural and conceptual equivalence in the acquisition and use of emotion words in a second language. *The Mental Lexicon* 3, 92-121.
- Segalowitz, N., Trofimovich, P., Gatlinton, E. & Sokolovskaya, A. (2008). Feeling affect in a second language: The role of word recognition automaticity, *The Mental Lexicon* 3, 47-71.
- Tsai, J. L., Simeonova, D. I. & Watanabe, J. T. (2004). Somatic and social: Chinese Americans talk about emotion. *Personality and Social Psychology Bulletin* 30(9), 1226-1238.

- Wierzbicka, A. (1999). *Emotions across languages and cultures: Diversity and universals*. Cambridge: Cambridge University Press.
- Xu, J. (2013). 'Zhongguo xuexizhe yingyu koutou xushi zhong de huayu pinjia yanjiu' [Investigating the appraisal system in spoken English narratives by Chinese EFL learners], *Waiyu Jiaoxue Yu Yanjiu* [Foreign Language Teaching and Research] 45(1), 69-79.
- Xu, J. & Jia, Y. (2011). BFSU Qualitative Coder 1.1. National Research Centre for Foreign Language Education, Beijing Foreign Studies University.
- Xu, J. & Jia, Y. (2011). BFSU Qualitative Explorer 1.1. National Research Centre for Foreign Language Education, Beijing Foreign Studies University.

*Vivian Lee
Department of Linguistics
SOAS, University of London
Thornhaugh Street, Russell Square,
London WC1H 0XG
England*

Vivian_Lee@soas.ac.uk

*Lu Lu
Department of Linguistics
SOAS, University of London
Thornhaugh Street, Russell Square,
London WC1H 0XG
England*

lu_lu@soas.ac.uk

WORD FORMATION IN PHASE THEORY

MAKIKO MUKAI
(University of Kochi)

Abstract

Recursive compounds, such as *peanut butter sandwich* pose a problem within Phase Theory (Chomsky 2001, 2008, Marantz 2007) with regard to Economy Condition. Harley (2009) uses Baker's (1988) incorporation, arguing compounding shows a case of morphology-as-syntax, as it is productive. But if roots are involved in word formation, there is no feature to trigger the incorporation. So her theory cannot be right. I propose a solution to this problem and show how this structure is applicable to recursive compounds.

1. Introduction

Recursion is a fundamental property of human language that potentially differentiates language both from other human cognitive domains and known communication systems in animals (Hauser, Chomsky, Fitch 2002, Corballis 2011). The main aim of this paper is to propose structure for recursive compounds (Roeper, Snyder, Hiramatsu 2002, Bisetto 2010, Tokizaki 2011) within Phase Theory (Chomsky 2001, 2008, Marantz 2007). Therefore, it is not to criticize other frameworks of word-formation, but to understand some aspect of recursion in Phase Theory.

Before going on to talk about recursion in compounding, let us first define what recursion in language is. Summarising a number of linguists, such as Chomsky (1965), Ralli (2013), (Bisetto 2010), Corballis (2011), and many others, I defined recursion as follows in Mukai (2013): it is a phenomenon of embedding structures within structures in cyclic fashion to create sentences, as complex and long as we like. Here, complex means embedding of phrases within phrases of the same kind. In principle it is possible to construct a limitless embedding structures in human language, at least, within the limitless of one's memory and processing capacity. According to Miyagawa & Nóbrega (2015) merge is the recursive operation of the language faculty. In this paper I follow this claim and use Phase Theory to propose structure for recursive compounds.

In this paper I define recursive compounds as compounds with embedding structures within structures. Bisetto (2010) and Tokizaki (2011) define recursive compounds with head addition, not without head addition. Let us show the difference between them.

- (1) Waste disposal plan
- (2) American English teacher

In (1), the compound *waste disposal* is expanded on the left side by means of the merger of a new constituent, *plan*. In contrast, in (2) the compound *English teacher* is not expanded by merging another head, but it is expanded by merging a modifier *American*. Compounds such as (2) are called iterative compounds by Bisetto (2010). However in this paper I will not follow their definition but argue both (1) and (2) types are recursive compounds in English, Mainland Scandinavian and Japanese and some other languages in the world.

The type (1) is left-branching compound while (2) is right-branching compound. The difference in the branching corresponds to a meaning difference. For example, the example

waste disposal plan means 'plan for waste disposal', not '*disposal plan for waste*'. In contrast, *American English teacher* means 'English teacher from America'.

This paper will be organised as follows. First, characteristics of recursive compounds will be explained. This section is followed by another section where I explain and criticize Harley's (2009) theory of compound word formation in Distributed Morphology. Then based on the criticisms I will propose a new theory in Phase Theory for word-formation. The conclusion of this paper will be the new analysis and implications for future research.

2. Characteristics of Recursive Compounds

2.1. Recursive compounds in different languages

As discussed in the previous section recursion can be seen at word level just as at phrase formation. Also, recursive compounds, both right-branching and left-branching is often accepted in many languages of the world, including Japanese, English, Mainland Scandinavian, Chinese, Korean, Dutch, German, Hungarian, Finnish, Thai and Vietnamese.

The first two examples are right-branching and the last two ones are left-branching recursive compounds, except in Korean, Vietnamese and Thai where only one type of branching can be found.

- | | | |
|-----|--|----------|
| (3) | [kagaku [benkyo beya]]
science study room | Japanese |
|-----|--|----------|

The word *room* in Japanese is translated as *heya* and when this word is the second constituent of a compound, the initial sound /h/ becomes voiced. So the word *benkyo beya* clearly shows it is a compound word.

- | | | |
|------|---|-------------------------------|
| (4) | [kochi [kenritsu daigaku]]
Kochi prefectural university
'University of Kochi' | Japanese |
| (5) | [booeiki gaisha] shachoo]
trade company president | Japanese |
| (6) | [[nise danuki] shiru]
mock badger soup | Japanese |
| (7) | [mail [delivery service]] | English |
| (8) | [student [feedback system]] | English |
| (9) | [[waste disposal] plan] | English |
| (10) | [[peanut butter] sandwich] | English |
| (11) | [barn-[bok-klub]]
child-book-club | Mainland Scandinavian |
| (12) | [aften-[komputer-klasse]]
evening-computer-class | Mainland Scandinavian |
| (13) | [[jule-mand] costume]
Christmas-man costume
'Santa Clause costume' | Mainland Scandinavian |
| (14) | [[bo-stad-s] kvarter]
live-place-Link area
'residential area' | Mainland Scandinavian |
| (15) | [informatie-[terugkoppeling-system]]
information-feed back-system | Dutch
(Google Netherlands) |

- (16) [beroep-s-[auto+andelaar]
professional-Link-car+delar
'professional car dealer' Dutch
- (17) [[grond-wet-s]-artikel]
ground-law-Link-article
'constitution' Dutch
- (18) [[scheep-s-bouw] maatschappij]
ship-building-Link+company Dutch
- (19) [Plastik-[garten-zwerg]]
plastic-garden-dwarf
'plastic garden dwarf' German
- (20) [Nord-[bahn-hof]]
North-train-court
'North station' German
- (21) [[fahrrad-kurier] jacket]
bicycle-courier jacket
'bicycle courier jacket' German
- (22) [[Blut-[druck]-apparat]
blood pressure apparatus German
- (23) [Barbie [selk ä[reppu]]
Barbie rucksack
(Google Finland) Finnish
- (24) [sein ä[kirja-hylly]]
wall book case
(Google Finland) Finnish
- (25) [[huone-kalu]-tehdas]
room+thing+factory
'furniture factory' Finnish
- (26) [[ti-tarve]-myynti]
home+need+scale
'household sale' Finnish
- (27) [fal-[könyv-szekr ény]
wall-book-case
(Google Hungarian) Hungarian
- (28) [fedett-[aut ó[parcol ó]]
roofed-car-park
'carport'
(Hungarian Google dictionary) Hungarian
- (29) [[v ér-nyom às]-mérő]
blood pressure apparatus
(Kiefer 2009: 527) Hungarian
- (30) [[p ó kerek] csapagy]
spare wheel bearing
(Tokizaki 2013: 290) Hungarian
- (31) [suān-[zh òng-d ú]]
acid-hit-position
'acidosis'
(Packard 2004:178) Chinese
- (32) [tiě-[f àn-[wǎn]]
iron-rice-bowl
'secure job'
(Packard 2004:178) Chinese
- (33) [[f èw ù chǔ zhì] j hu à]
waste disposal plan
(Tokizaki 2011:9) Chinese

- | | | |
|------|---|--|
| (34) | [[zhongguo gongchan] dang]
China communist party
'Communist Party of China' | Chinese |
| (35) | [upyeon-[baedal-seobiseu]]
mail-delivery service | Korean (right-branching)
(Google Koera) |
| (36) | [on-chən]-yok]
warm-spring-bathe
'bathing in a hot spring' | Korean (left-branching) |
| (38) | [[duw̃ŋ cʌn] [x̃i]
line foot sky
'horizon' | Vietnamese |
| (37) | [[b ɣ̃ sɛ] b̃]
wheel vehicle cow/ox
'cow/ox cart wheel' | (Dang 2013: 21) |
| (38) | [[raan khai] khanom]
shop sell sweets
'candy store' | Thai

Post (personal communication) |

As the examples show recursive left-branching and right-branching compounds are productive in these languages. The degree of productivity is different in different languages, but these languages seem to have recursive compounds.

It should be noted that Tokizaki (2013) comments that Thai does not have productive recursive compounds and three-member-compounds are found in menu of dishes. However, Post (personal communication) argues that there are three-member-compounds in Thai, too. Certainly there needs to be more research on this language. However, the definition for 'productivity' of word-formation seems to conflict among different researchers. I leave this to future research.

Also, even though they do not allow as productive recursive compounds as the languages above, Italian and French, allow exocentric VNN compounds. They are right-branching compounds (iterative in Bisetto (2010)) and the VN compound becomes the complement of a verb. Even though they are not as productive as NNN recursive compounds in the above languages, recursion and the presence of self-embedding in such constructions, show that these compounds are derived by a Merge (Miyagawa & Nóbrega 2015). Examples of this type of compounds from these languages are as follows.

- | | | |
|------|--|-------------------------------|
| (39) | [porta-[stuzzica-denti]]
carry 3.pres-pick-teeth
'toothpick-holder' | Italian

Bisetto (2010) |
| (40) | [porta-[ombrello bici]]
carry 3. pres-umbrella bike
'umbrella holder for bike' | Italian (Google Italia) |
| (41) | [porte-[parapluie v ɛ̃o]]
carry 3. pres-umbrella bike
'umbrella holder for bike' | French (Google France) |

2.2. Characteristics of recursive compounds in Japanese, English and Mainland Scandinavian

The last sub-section showed some examples of recursive compounds in different languages of the world. The aim of this paper is to propose structures for recursive

compounds which are appropriate for any language of the world. As time and space are limited to show characteristics of recursive compounds in all of the languages discussed in 2.1, I focus on Japanese English and Mainland Scandinavian.

Recursive compounds are phrase-like in the following sense. First, they are pronounced like two independent words and there is a slight pause in between the second and third constituents. Also, they are very productive and internal structure is visible to syntax. However, they do show word-like characteristics in that functional category is excluded and it is impossible to delete part of them.

Let us show the above arguments are true. First the examples from Japanese (3)-(6) are shown.

- (42) [kagaku|[benkyo beya]]
- (43) [Kochi|[kenritsu daigaku]]
- (44) [[nise danuki]| shiru]
- (45) [[booeiki gaisha]| shachoo]

‘|’ stands for a short pause. In both right-branching and left-branching compounds there is a short pause. However, it is impossible to have a functional category.

- (46) [kagaku [benkyo (*no) beya]]
science study (*GEN) room
- (47) [kochi [kenritsu (*no) daigaku]]
- (48) [[nise (*no) danuki] shiru]
mock (*GEN) badger soup
- (49) [[booeiki (*no) gaisha] shachoo]
trading (*GEN) company president

The example (49) is possible if there is a short pause between the words *kenritsu* and *daigaku*. However when the two words are pronounced together like a compound word a functional category cannot intervene.

Also, it is impossible to delete a part of the compounds.

- (50) *Gakko de [kagaku [benkyo ~~beya~~]- \emptyset tsuka-tta-ga, daigaku de-wa
School dat [science [study ~~room~~]-ACC use-PAST but, university dat-const
[kagaku [benkyo beya]] o tsuka-wana-ka-tta.
[science [study room]] ACC use-NEG-past
‘I used science study room at school but I did not at university.’
- (51) *Ibaragi de [[nis danuki] ~~shiru~~]- \emptyset tabe-ta ga, Kochi de [[nise danuki]
Ibaragi dat [[mockbadger] ~~soup~~]-acc eat- PAST but, Kochi dat [[mock badger]
shiru o tabe-naka-tta.
eat-neg-PAST
‘I ate mock badger soup in Ibaragi, but didn’t eat in Kochi.’
- (52) *Kobe de wa [[booeiki gaisha] ~~shachoo~~]- \emptyset ni a-tta ga, Osaka de wa
Kobe dat top [[tradingcompany]pres.—DAT meet-PAST but, Osaka DAT TOP
[[booeiki gaisha]shachoo] ni awa-na-ka-tta.
trading company pres. DAT meet-NEG-PAST
‘I met a trading company president in Kochi but I did not in Osaka.’

The functional category, genitive case marker cannot intrude the examples or it is impossible to delete its part in recursive compounds in English and Mainland Scandinavian, like in Japanese. First let me show that there is a short pause between the constituents.

- (53) [mail|[delivery service]]
- (54) [student| [feedback system]]
- (55) [[peanut butter]| sandwich]
- (56) [[waste disposal]|plan]
- (57) [barn|[bog-klub]]
child book club
- (58) [aften|[komputer-klasse]]
Evening computer class
- (59) [jule-mand]| kostume]
Christmas-man costume
'Santa Claus man'
- (60) [[bo-stad]-s|-kvarter]
Live-place-link-area

Right-branching has a short pause after the first constituent, which is like a modifier of the 'compound word', while left-branching has one after the second constituent. This phenomenon is also seen in Japanese, like the examples (44)-(47) show. How about intervention of functional element?

- (61) [mail [delivery (*of) service]]
- (62) [student [feedback (*of) system]]
- (63) [barn [bog-(*s)-klub]]
child book-(*link)-club
- (64) [aften [komputer-(*s)-klasse]]
evening computer-(*link)-class
- (65) [peanut (*of) butter] sandwich]
- (66) [[jule-(*s)-mand] kostume]
- (67) [[bo-(*s)-stad]-s-kvarter]

The above examples show that both right-branching and left-branching recursive compounds cannot have a functional element in-between.

Also, it should be noted here that there is a linking morpheme in Scandinavian languages (also found in other non-English Germanic languages). According to Josefsson (1997), it is a morpheme without independent meaning. The linking morpheme is used to get this structure and interpretation. Without the linking morpheme, Swedish is strictly right-branching compounds. For example, without a linking morpheme, the compound in (69) is interpreted as "place area for resident", which cannot be logical in the real world. Bisetto (2010) and Mukai (2013) discuss linking morphemes found in compounds in other languages of the world. However there does not seem to have any consensus on their function.

In Section 2 it has been shown that there are examples of endocentric right-branching and left-branching recursive compounds in different languages and exocentric right-branching compounds in Romance languages. In addition focusing on recursive compounds in Japanese, English and Mainland Scandinavian I have shown both phrase-like and lexical characteristics of these compounds.

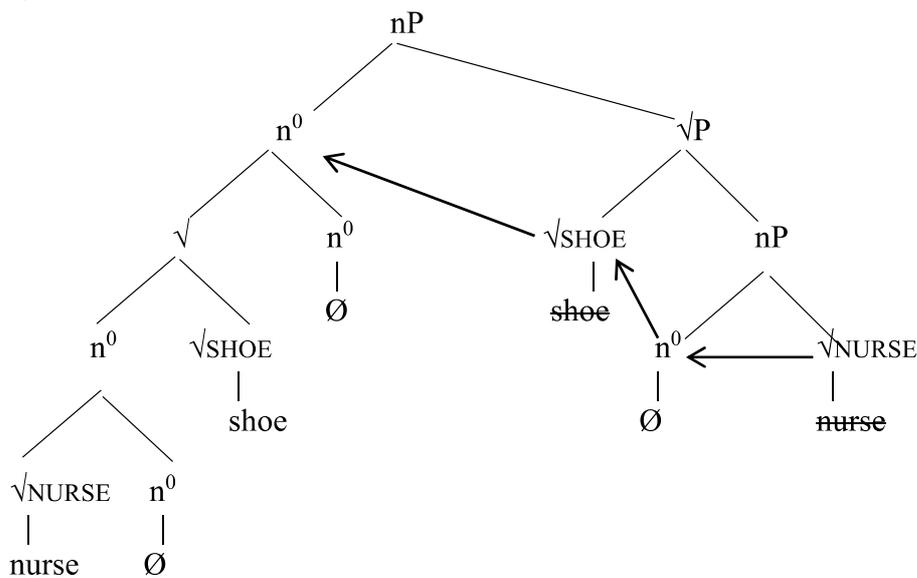
The next section discusses and criticizes Harley's theory of word-formation.

3. Harley's word-formation

In Distributed Morphology it is argued that all generalised transformations are assumed to be of the same kind, showing that they take place in a single module (Marantz 2007). Thus, there is no single place called a 'lexicon' distinct from syntax where some generalised transformations take place. Another point is that only formative aspects of morphemes, but not phonological aspects, are contained in the list of morphemes. With above in mind Harley (2009) argues compounding shows a case of morphology-as-syntax as it is productive and uses Baker's (1988) incorporation for derivation of words.

The following is a structure for primary compounds *nurse shoe*, proposed by Harley.

(68)

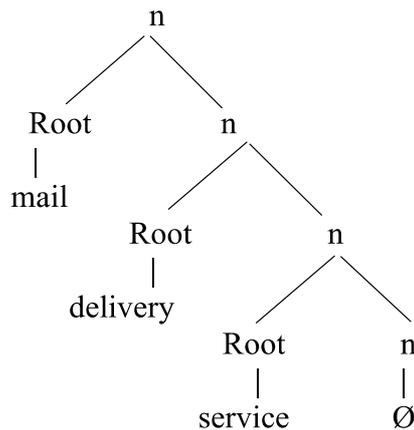


Harley (2009: 140)

Harley uses Baker's (1988) incorporation to capture the Lexical Integrity, namely indivisibility of the elements in the compounds, the impossibility of the movement out of them, and the impossibility of discourse antecedent from within a compound. The structure shows that the roots without categorical feature (Marantz 2007) are merged with their category-determining feature bundle, or otherwise, they cannot appear in the narrow syntax. Also, Harley assumes the modificational relationship between the non-head (*nurse*) and the head (*shoe*) in the compound.

Harley does not propose a structure for recursive compounds, discussed in Section 2. However, I assume that the structure is as follows.

(70)



The above structure is one for right-branching recursive compounds. This structure is constructed as follows: First, a root without word class feature (Zhang 2007) is merged with a syntactic head, thus turning the root into an n (Marantz 2007). This is labelling, in terms of Chomsky (2008). Then, another root is merged to form compound word and this ‘compound’ is transferred to the interpretational component and spelled out as phase (Chomsky 2001, 2008).

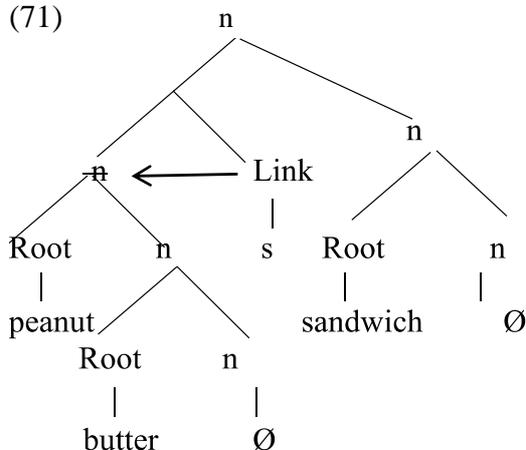
The reason why the two-member compound is assumed to be a phase is that by assuming it is a phase we can capture the Lexical Integrity of compounding (Section 2). That is, as once the compound is a phase, it is impossible to delete part of it or no functional element can intervene between them.

Another root is merged to form the right-branching compound and the head of the whole compound is the ‘right-most’ element.

The three roots are not merged immediately because the interpretation of the whole compound allows for alternative and compositional interpretation. Also, roots cannot be directly combined with each other as they are void of grammatical feature (Zhang 2007, Mukai 2006, 2008). As a result the n is the head of the whole compound word, as there is no categorizing feature to be interpreted semantically and phonologically.

How about left-branching recursive compounds? I propose the following structure for this type of compounds.

(71)



Like the structure for right-branching recursive compounds, the structure for left-branching recursive compounds does not need any movement, unlike in Harley’s theory. The

structure above is derived in the following ways: First, a root without word class feature is merged with a syntactic head, thus turning the root into an *n* (Marantz 2007). Another root is merged to form a two-member compound word. Here, the compound is transferred to the interpretational component and spelled out. The two-member compound with its syntactic head is interpreted semantically and phonologically. The interpretation is different depending on the roots.

Like in the structure (72) the two roots are not merged immediately in (73), as the interpretation of the whole compound allows for alternative and compositional interpretation. Also, the reason why the two-member compound is assumed to be a phase is that by assuming it is a phase we can capture the Lexical Integrity of compounding (see Section 2). That is, as once the compound is a phase, it is impossible to delete part of it and to be referred back as a discourse antecedent.

The categorizing head is merged, as in recursive compounds there is a slight pause between the second and third constituent of the recursive compound (Tokizaki 2011, 2013). Here, Tokizaki argues that there is a head.

However, as it is impossible to have two heads in the syntax, a linking morpheme, which is phonetically there in Mainland Scandinavian and several other languages (see Section 2), but not in other languages, like Japanese or English, is merged to check the categorizing head. The resulting structure is merged with another root which is merged with a categorizing head, which is derived in parallel. As a result, the categorizing head *n* is the head of the whole compound.

Unlike in Harley's theory there is no unnecessary movement in this theory.

In this section I proposed structures for right-branching recursive compounds and left-branching recursive compounds in Phase Theory. Unlike Harley's theory there is no movement but only the operation merge. Assuming a phase in the structure can capture the Lexical Integrity of compounds, as discussed in Section 2. The linking element is there for Economy reason to check the feature on the head.

The present theory is applicable for recursive compounds in the languages discussed in Section 2.

What about productivity of recursive compounds? As mentioned in Section 2.1 it is argued that Romance languages do not allow as productive recursive compounds as in the other languages, especially Germanic languages. Tokizaki (2013) argues that this might be due to phonology. I leave this problem for future research.

5. Conclusion

In this paper I proposed structures for recursive compounds in Phase Theory. As discussed in Section 2 recursive compounds show both phrase-like and word-like characteristics. By assuming there is a phase in the structure Lexical Integrity of recursive compounds can be captured without any unnecessary movement, unlike in Harley's theory.

However, there remains some problems to be solved in this theory. First, what kind of feature is there in the head? How is structure for right-branching recursive compounds with a linking morpheme inside in German or Greek languages? More research is required to answer these questions, along with many other questions.

References

Baker, M. C. (1998). *Incorporation*. Chicago & London: Chicago University Press.

- Bauke, L. (2009). *Nominal Root Compounds*. Ph.D dissertation, UMass Amherst.
- Bisetto, A. (2010). Recursiveness and Italian Compounds. *SKASE Journal of Theoretical Linguistics* 7, 14-35.
- Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Chomsky, N. (2001). Beyond Explanatory Adequacy. *MIT Occasional Papers in Linguistics* 20, 1-28.
- Chomsky, N. (2008). On Phases. In Freidin, R., Otero, P., & Zubizarreta, M. L. (eds.), *Foundational Issues in Linguistic Theory. Essays in Honor of Jean-Roger Vergnaud*, 133-166, Cambridge, MA: MIT Press.
- Corballis, M. C. (2012). *The Recursive Mind: The Origins of Human Language, Thought, and Civilization*. Princeton: Princeton University Press.
- Dang, P. H. (2013). *Does prosodic word recursion cause phonetic initial strengthening?* MA Thesis, California State University, Fresno.
- Harley, H. (2009). Compounding in Distributed Morphology. In Lieber, R., & Stekauer, P. (eds.), *The Oxford Handbook of Compounding*, 129-144. Oxford: Oxford University Press.
- Hauser, M. D. Chomsky, N., & Fitch, T. (2002). The Faculty of Language: What is it, Who has it, and How did it Evolve? *Science* 98, 1569-1579.
- Josefsson, G. (1997). *On the Principles of Word Formation in Swedish*. Lund: Lund University Press.
- Kiefer, F. (2009). Uralic, Finno-Ugric: Hungarian. In Lieber, R., & Stekauer, P. (eds.), *The Oxford Handbook of Compounding*, 527-541. Oxford: Oxford University Press.
- Marantz, A. (2007). Phases and words. In Choe, S.H. (ed.), *Phases in the Theory of Grammar* (pp. 191-222). Seoul: Dong-In Publishing Co.
- Miyagawa, S., & Nóbrega, N.A. (2015). The precedence of syntax in the rapid emergence of human language in evolution as defined by the integration hypothesis. *Frontiers. Psychology*, 6: 271.
- Mukai, M. (2006). *A Comparative Study of Compound Words in English, Japanese and Mainland Scandinavian*, Ph.D dissertation, University of Newcastle.
- Mukai, M. (2008). Recursive Compounds. *Word Structure* 1 (2): 178-198.
- Mukai, M. (2013). Recursive Compounds and Linking Morpheme. *International Journal of English Linguistics* 3 (4), 36-49.
- Packard, J. (2004). *The Morphology of Chinese: A Linguistic and Cognitive Approach*. Cambridge: Cambridge University Press.
- Ralli, A. (2013). Compounds and its Locus of Realization: Evidence from Greek and Turkish. *Word Structure* 6 (2), 181-200.
- Roeper, T., Snyder W., & Hiramatsu, K. (2002). Learnability in a Minimalist Framework: Root Compounds, Merger, and the Syntax-Morphology Interface, 25-37. In Lasser, I. (ed.). *The Process of Language Acquisition*. Frankfurt: Peter Lang Verlag.
- Tokizaki, H. (2011). The Nature of Linear Information in the Morphosyntax-PF Interface. *English Linguistics* 28 (2), 227- 257.
- Zhang, N. (2007). Root Merger in Chinese Compounds. *Studia Linguistica* 61 (2), 170- 184.

*Makiko Mukai
Department of Cultural Studies
2-22 Eikokuji
Kochi 780-8515
Japan*

mukai@cc.u-kochi.ac.jp

FROM “HEN” TO ADJECTIVAL MODIFICATION IN MANDARIN CHINESE

FANGFANG NIU

(Queen Mary University of London)

Abstract

This paper explores the question of why adjectives cannot function as predicates by themselves in Mandarin Chinese (MC). Instead, degree morphemes, question particles and other elements are required to co-occur with adjectives. Following Rooth’s (1992) and Ramchand’s (1996) discussion on focus interpretation, I propose that the function of these elements is to create a set of alternative propositions, which is needed to satisfy the [+FOC] feature of the Pred head.

1. Bare adjectives are highly restricted in MC

It is an important characteristic of MC syntax that when acting as predicates, adjectives are normally accompanied by degree morphology (Sybesma 1999, Dong 2005, Huang 2006, Grano 2008, Liu 2010, Grano 2011, Zhang 2015, among others). A very common one of these degree morphemes is the adverb *hen* ‘very’:

- (1) a. ??Zhangsan gao.
Zhangsan tall
- b. Zhangsan hen gao.
Zhangsan very tall
‘Zhangsan is (very) tall.’

Without the appearance of *hen* ‘very’, sentence (1a) is unacceptable under the meaning ‘Zhangsan is (very) tall’. Apart from *hen*, degree complements such as *budeliào* ‘incredibly’ can also save sentence (1a):

- (2) Zhangsan gao de¹ *(budeliào).
Zhangsan tall DE incredibly
‘Zhangsan is incredibly tall.’

In addition to degree morphology, a variety of elements are available to rescue sentence (1a). This includes *bi* comparative phrases, question markers, negators, quantity phrases, among many others. In the following, I will give an example for each of these cases.

The *bi* phrase *bi Lisi* can appear before the adjective *gao* to form a comparative construction.

- (3) Zhangsan bi Lisi gao.
Zhangsan BI Lisi tall
‘Zhangsan is taller than Lisi.’

¹ Here, *de* is a complement marker.

The adjective *gao* ‘tall’ can be followed by the question marker *ma* to form a yes-no question.

- (4) Zhangsan *gao ma*?
Zhangsan tall Q
‘Is Zhangsan tall?’

Similarly, *gao* ‘tall’ can form an A-not-A (yes-or-no) question.

- (5) Zhangsan *gao bu gao*?
Zhangsan tall not tall
‘Is Zhangsan tall?’

Quantity phrases (QP), which are composed of numerals and units of measure, can appear before the adjective as well:

- (6) Zhangsan *liang mi gao*.
Zhangsan two meter tall
‘Zhangsan is two-meter tall.’

What is more, (1a) becomes fine when it is negated:

- (7) Zhangsan *bu gao*.
Zhangsan not tall
‘Zhangsan is not tall.’

In the above, I illustrated a few of the range of elements that can appear with the adjective *gao* ‘tall’ in (1a) to form a well-formed subject-predicate sentence. The main concern of this paper is why bare adjectives cannot appear on their own and how these different elements turn the adjective into a legal predicate. This will be discussed in section 2 and 3. Before I move on, I would like to point out that actually, sentence (1a) is possible under a comparative reading: Zhangsan is taller than some person/people in the context. What is worth noting is that in the case, Zhangsan is necessarily focused. This is an important clue to the analysis I am going to propose, so I will focus on discussing this issue in the next section.

2. Cases where adjectives do stand on its own

There are only a few cases where adjectives do stand on its own. All these cases happen in a contrastive situation where one entity/feature is contrastively focused. For instance, in the following sentence, the *wh*-word *shui* ‘who’ is stressed.

- (8) Zhangsan he Lisi, *shui_F² gao*?
Zhangsan and Lisi who tall
‘Zhangsan and Lisi, who is taller?’

(8) is a *wh*-question sentence, in which the predicate is a bare adjective *gao* ‘tall’. The subject *shui* ‘who’ is stressed and the whole sentence carries a comparative reading, as

²A small F is put after *shui* to indicate that it is stressed phonologically.

indicated by the translation. As an answer to question (8), the following sentence is possible, in which the subject *Zhangsan* is stressed:

- (9) Zhangsan_F gao.
 Zhangsan tall
 ‘Zhangsan is taller.’

Another case is when the adjective is stressed, for example, when answering the yes-no question (4) or (5), the sentence (10) is fine.

- (10) Zhangsan gao_F.
 Zhangsan tall
 ‘Zhangsan is tall.’

Phonologically, the adjective *gao* ‘tall’ is pronounced longer and stronger. Semantically, it has the flavor of confirming the fact that Zhangsan is tall, rather than short.

Another contrastive example is given below, in which both the adjectives *gao* ‘tall’ and *ai* ‘short’ are stressed. This sentence could also be interpreted as a contrast between the adjective ‘tall’ and its antonym ‘short’.

- (11) Zhangsan gao_F, Lisi ai_F.
 Zhangsan tall Lisi short
 ‘Zhangsan is tall while Lisi is short.’

In brief, in (8) and (9), what is in contrast is the subject, i.e., Zhangsan or Lisi, as indicated by the focus/stress marker F, whereas in (10) and (11), it is the feature ‘tall’ or ‘short’ that is contrasted. These facts suggest that the way adjectives are introduced as predicates in MC is closely related to focus interpretation.

3. The Pred_[+FOC] analysis

3.1. Focus interpretation & alternative semantics

Rooth (1992) and Ramchand (1996) argue that the notion of a set of alternatives is widespread across languages. Focus is an important mechanism of creating alternative semantics and the alternative semantics of a sentence is a set of alternative propositions created by making substitutions in the position of the focused phrase:

- (12) $[[[s[Mary]^F \text{ likes Sue}]]]^f = \{ \text{like}(x, s) \mid x \in E \}$

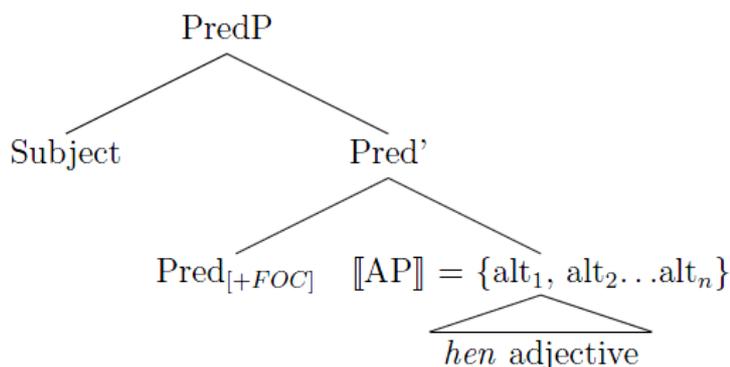
E represents the domain of individuals. ‘Mary’ is focused, the alternative semantics of ‘Mary likes Sue’ is the set of propositions created by substituting Mary, i.e., $\{ \text{like}(x, s) \mid x \in E \}$. According to Rooth (1992), a range of linguistic elements are sensitive to alternative semantics signaled by focus, one of them is the English adverb *only*:

- (13) a. Mary only introduced Bill to [Sue]_F.
 b. Mary only introduced [Bill]_F to [Sue].

Sentence (13b) is untrue in a scenario that Mary introduced both Bill and Tom to Sue.

So my claim is that in MC, bare adjectives are not predicative in nature, and they need to be turned into predicates. Following Rooth and Ramchand's idea, I propose that Mandarin adjectives are introduced as predicates by creating alternative propositions. Specifically, in predication constructions, morphemes such as *hen* and *ma* perform the function of generating alternative propositions by building contrastive pairs. Structurally, I will argue that PredP is projected in Mandarin predication constructions (Adger & Ramchand 2003) and the Pred carries a [+FOC] feature. This proposal can be formalized as the following:

(14)



The [+FOC] feature of the Pred head needs to be satisfied by a set of alternatives, therefore, elements such as *hen*, negators, question particles, etc., are required to create alternative propositions. As an illustration, in the following sentence, the *wh*-word *na* 'which' denotes the alternative semantics, which is the group of students known in the context, it could be {Zhangsan, Lisi, Wangwu...}.

(15) Na ge xuesheng (hen) gao?
 Which CL student very tall
 'Which student is (very) tall?'

The morpheme *hen* is optional in this case. It is worth noting that it is a standard view that there are two *hen* in MC: one has syntactic functions, which is the one discussed in this paper, while the other one purely indicates degree, and is normally stressed (Li & Thompson 1981, Chui 2000, etc.), such as the one above: *na* 'which' provides the set of alternatives, *hen* is a degree intensifier.

Under this analysis, cases where the subject or the adjective is focused follows naturally: to create a set of alternatives by building contrastive scenarios. For instance, for sentence (9), the set of alternatives could be {Zhangsan, Lisi}; for sentence (10), {tall, not tall}; and for (11), {taller than Lisi, as tall as Lisi, shorter than Lisi}.

3.2. The application of the Pred_[+FOC] analysis

In the next, I will go through those cases where there is no focus intonation and elements such as *hen* and *ma* co-occur with the adjective, to show how the current analysis captures these data. First of all, when the adjective is accompanied by *hen* 'very':

(16) Zhangsan hen gao.
 Zhangsan very tall
 'Zhangsan is (very) tall.'

hen indicates a set of degrees such as {extremely, very, moderately, a bit}. The above sentence can be interpreted as Zhangsan is very tall, but not moderately tall or extremely tall.

More obviously, in the following *bi* comparative sentence, the *bi* phrase *bi Lisi* indicates a set of alternatives of the height difference between Zhangsan and Lisi (the stand of comparison): {Zhangsan is taller than Lisi, Lisi is taller than Zhangsan, Zhangsan is as tall as Lisi}.

- (17) a. Zhangsan *bi* Lisi gao.
Zhangsan BI Lisi tall
'Zhangsan is taller than Lisi.'
- b. *Zhangsan *bi* Lisi *hen* gao.
Zhangsan BI Lisi very tall

Even more interestingly, when *bi Lisi* is present, the degree adverb *hen* is not allowed (17b). This supports the assumption that the *bi* phrase and *hen* has the same function and therefore there is no need for them to show up at the same time in the same sentence. The following two types of yes-no questions can be analyzed in a similar way.

- (18) Zhangsan gao ma?
Zhangsan tall Q
'Is Zhangsan tall?'
- (19) Zhangsan gao bu gao?
Zhangsan tall NEG tall
'Is Zhangsan tall?'

According to Liu (2010), there exists the degree value of Zhangsan's height and the contextually determined standard degree of human height and it is the relation of these two degrees that is being asked about in the yes-no question. Alternatively, this above two sentence can be interpreted as Zhangsan's height > the standard human height or Zhangsan's height < the standard human height. In these cases, it is the question particle *ma* and the A-not-A question form that check the [+FOC] feature of the Pred head. As for the negation case, following Rooth (1992), Lee (2001) proposes that *bu* is a focus sensitive operator which introduces a set of alternatives to the part that is negated. Specifically, in (20) below, there is an alternative to 'Zhangsan is not tall', which is 'Zhangsan is tall'.

- (20) Zhangsan bu gao.
Zhangsan NEG tall
'Zhangsan is not tall.'

In fact, the A-not-A question in (19) can be understood in this way as well. The set of alternatives is {Zhangsan is tall, Zhangsan is not tall}.

With respect to the following case where the adjective is accompanied by a quantity phrase, it can be said that it creates a set of different quantity values (heights) {1.5 meters, 1.8 meters, 2 meters...}.

- (21) Zhangsan liang-mi gao.
Zhangsan two-meter tall
'Zhangsan is two-meter tall.'

The sentence below is very interesting, where the adjective co-occur with the sentence final particle *le*, denoting a change of state:

- (22) Zhangsan gao le.
Zhangsan tall LE
'Zhangsan got taller.'

Since this sentence denotes change from one state to another, it is very plausible to assume that the set of alternatives include the different states of the subject: {1.5 meters tall, 1.6 meters tall, 1.7 meters tall}.

As shown by the examples above, where elements such as degree morphemes, question particles, negators appear, there is indeed a set of alternatives present. It can be concluded that the function of those elements is creating alternative semantics. And this explains why they are obligatory in adjective predicates in MC: to check the [+FOC] feature of the Pred head.

4. Further evidence

To restate the proposal again, the Pred head in MC carries a [+FOC] feature and this feature needs to be satisfied by a set of alternatives. Elements such as degree morphemes, question particles, negators are required to create alternative semantics. It can be seen that the alternative/contrastive semantics plays an important role in this process. Therefore, it can be predicted that when contrastive semantics cannot be built, bare adjectives must be banned.

4.1. When there is no contrast

As discussed above, when there is no degree morpheme or any other elements accompanying the adjective, in a well-formed sentence such as the following, the two adjectives must form a contrast, to compose a set of alternatives.

- (23) Zhangsan gao_F, Lisi ai_F.
Zhangsan tall Lisi short
'Zhangsan is tall while Lisi is short.'

However, as pointed out by Dong (2005), unlike (23), the following sentence is unacceptable.

- (24) Zhangsan gao, Lisi gaoxing.
Zhangsan tall Lisi happy
'Zhangsan is tall and Lisi is happy.'

The two adjectives *gao* 'tall' and *gaoxing* 'happy' do not form a contrast with each other, as a result, the bare adjectives cannot act as the predicates, consequently, sentence (24) is bad.

Before turning to the next part, I would like to mention that Grano (2008) claims that if embedded, clauses with bare adjectives could be acceptable. The example he gives is below:

- (25) Wo zhidao [Zhangsan gao], dan mei xiangdao ta zheme gao.
I know Zhangsan tall, but NEG.PRF expect he this tall
'I knew Zhangsan was tall, but I didn't expect he was this tall.'

He argues that *hen* is only required in the matrix clause and in embedded clauses, adjectives can appear in predicate position without the presence of degree morphology. However, this is not true. For instance, if I keep the first half of the above sentence and change the second half to that in (26a), the sentence becomes unacceptable. It would be better to have *hen* ‘very’ before the adjective, as shown in (26b):

- (26) a. ?Wo zhidao [Zhangsan gao], suoyi rang ta qu da lanqiu.
 I know Zhangsan tall, so let him go play basketball (bb)
 ‘I know Zhangsan is very tall, so I let him play basketball.’
- b. Wo zhidao [Zhangsan hen gao], suoyi rang ta qu da lanqiu.
 I know Zhangsan very tall, so let him go play bb
 ‘I know Zhangsan is very tall, so I let him play basketball.’

The only difference between (25) and (26a) is the relation between the two sub-clauses: in the former, it is transitional; while in the latter, it is causal. That is to say, in causal relation, such as (26a), bare adjectival predicates are not legal in embedded clauses. This may suggest that what makes (25) grammatical is the transitional relationship between clauses, more specifically, the contrast between the speaker’s presupposed height of Zhangsan and his actual height. This supports my claim that when there is no contrast/alternative such as in (26a), bare adjectives are not permitted, while when there is a contrast, even it is clausal such as in (25), bare adjectives are acceptable. Whether it is a matrix clause or an embedded clause does not make a difference.

4.2. When there is no Pred

Another predication of the current proposal is that when the Pred head is absent, the degree elements and others should not appear as well. This is indeed the case.

4.2.1. Small clauses

The first environment where Pred is not projected is in small clauses. The sequence *Lisi ai* ‘Lisi short’ in (27) below is generally regarded as an instance of small clauses (Tang 1998).

- (27) Zhangsan xian Lisi ai.
 Zhangsan disfavor Lisi short
 ‘Zhangsan disfavors Lisi for being short.’

According to native speakers, this sentence is completely fine. This is captured by my assumption: since Pred head is not present in *Lisi ai*, there is no reason for the degree morpheme to show up. In fact, when the degree marker appears, the acceptability of the sentence decreases greatly, as shown below.

- (28) *Zhangsan xian Lisi hen ai.
 Zhangsan disfavor Lisi very short

Cases where degree elements do appear, they are just degree intensifiers.

- (29) Zhangsan xian Lisi tai ai.
 Zhangsan disfavor Lisi too short
 ‘Zhangsan disfavors Lisi for being too short.’

It is interesting that *tai* ‘too’ is acceptable in the above sentence but not *hen*. Semantically, *tai* is stronger than *hen* and it also carries the speaker’s dissatisfaction with the excessive degree denoted by the adjective. So I will assume that *tai* ‘too’ in the above sentence is a pure degree intensifier and does not have any influence on the syntax of the sentence (more evidence is needed for this argument).

4.2.2. Pre-nominal modification

Similar reasoning applies to the modification cases in (30) where PredP does not exist.

- (30)a. hong hua
 red flower
 ‘a red flower’ or ‘red flowers’
- b. xiao juzi
 small mandarin
 ‘a small mandarin’ or ‘small mandarins’

Generally speaking, *hen* is not required when adjectives modify nominals, as shown by examples above. However, interestingly, whenever *hen* shows up, *de* must co-occur, as can be seen from the contrast between (30) and (31)³.

- (31)a. hen hong *(de) hua
 very red DE flower
 ‘a flower that is very red’ or ‘flowers that are very red’
- b. hen gao *(de) nan haizi
 very tall DE male child
 ‘a boy that is very tall’ or ‘boys that are very tall’

This may be because *hen hong de* in (31a) is a clause and that *hen* is obligatory to check the [+FOC] feature of the Pred head. As a matter of fact, it has been argued that phrases such as (31a) and (31b) are derived from relative clauses and *de* in these cases is a clausal marker (Larson 2009).

Moreover, in phrases where *hen* is not present, the adjective is necessarily stressed, as shown in (32), *hong de hua* ‘red flowers’ is in contrast with *huang de hua* ‘yellow flowers’:

- (32)a. Wo xihuan hong_F de hua.
 I like red DE flower
 ‘I like red flowers (not yellow ones).’

³Grano (2008) argues that all the above prenominal adjectives are within relative clauses, therefore, it is fine for them to show up bare (he believes elements such as *hen* are only necessary in matrix clauses). However, the general assumption is that only the *de* cases in (31) are relative clauses and in the *de*-less cases in (30) adjectives are merged with the nominals directly (Sproat and Shih 1991, Paul 2006, among others). As a result, Grano’s argument does not hold.

- b. Hong_F de hua haokan.
 red DE flower beautiful
 ‘red flowers (not yellow ones) are beautiful.’

If we adopt Larson’s assumption that *de* is a relative clause marker, the fact that *hong* ‘red’ in (32) is stressed serves as a further argument for my Pred_[FOC] proposal.

It is worth mentioning that the reduplicative form of the adjective, also known as complex adjectives (CA), can act as predicates directly:

- (33) Zhangsan gao-gao-de.
 Zhangsan tall-tall-DE
 ‘Zhangsan is tall.’

Compared with simple adjectives (SAs), CAs represent an intensified degree and sentence such as (33) could be viewed as a contrast between different degrees of tallness.

4.3. Non-gradable adjectives

Until now, all the adjectives I discussed are gradable adjectives, the other type of adjectives in MC is non-gradable adjectives such as *dui* ‘right’, *zhen* ‘authentic’, etc.. One property of these adjectives is that they normally cannot be modified by degree morphemes, however, they usually appear in *shi...de* constructions⁴, as shown below:

- (35) a. ??Zhe ge huaping hen jia.
 this CL vase very fake
 b. Zhe ge huaping shi jia de.
 this CL vase SHI fake DE
 ‘This vase is fake.’

The *shi...de* construction is generally considered to be a focus construction in the literature. In (35b) above, the adjective *jia* ‘fake’ is focused to form a contrast with its potential antonym *zhen* ‘authentic’. The whole sentence is to emphasize that this vase is fake rather than genuine.

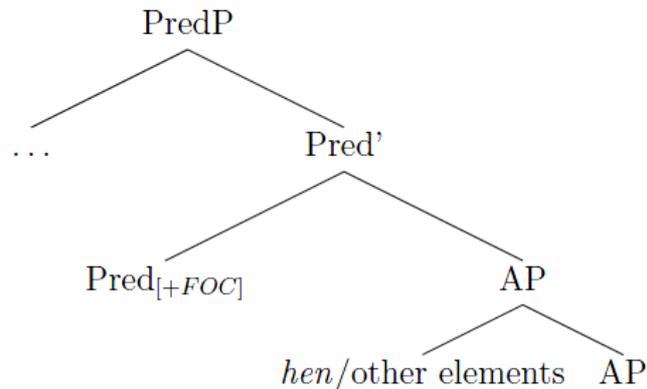
This fact above together with the discussion presented so far for gradable adjectives suggests that adjectives (both gradable and non-gradable) in MC are introduced as predicates by a functional projection PredP which bears an [+FOC] feature. The schema can be illustrated below:

⁴Gradable adjectives normally do not show up with *shi...de*:

- (34) ??Zhangsan shi gao de.
 Zhangsan SHI tall DE
 ‘It is true that Zhangsan is tall.’

This sentence is marginally acceptable.

(36)



Elements around the adjective, including stress intonation, degree expressions and other morphemes, perform the same role: creating a set of alternatives to satisfy the [+FOC] feature of the Pred head.

5. Previous analyses and their limitations

Generally speaking, there are three lines of research regarding the issue why bare adjectives are not allowed as predicates in MC.

5.1. The semantic type shifting approach

The semantic approach argues that adjectives are of a particular semantic type and they need the co-occurrence of other morphemes such as *hen* to convert them into the right type in predicate position, before combining with the subject (Huang 2006, Liu 2010, Zhang 2015). More specifically, it has been proposed that adjectives are of type *e* (Huang 2006) or $\langle d, \langle e, t \rangle \rangle$ (Liu 2010), and they require the co-occurrence of degree morphology such as *hen* (type $\langle e, \langle e, t \rangle \rangle$) or an operator POS (type *d*), respectively, to be converted into type $\langle e, t \rangle$ in predicate position. Following Kennedy (1997) and Liu (2010), Zhang (2015) also argues that *hen* is a POS marker based on the fact that it occurs in positive constructions only. It is of semantic type $\langle \langle d, \langle e, t \rangle \rangle, \langle e, t \rangle \rangle$ and *s*-selects gradable adjectives which are type $\langle d, \langle e, t \rangle \rangle$.

However, this line of analyses fails to capture the fact that degree morphology is not needed when adjectives modify nouns attributively, unless it assumes that the semantic type of adjectives changes in modification constructions such as the ones below:

- (37) xiao (de) juzi
Small DE orange
'small oranges'

Moreover, Liu, Huang and Zhang's analyses cannot explain why morphemes such as question marker *ma*, etc. can save sentences without the appearance of *hen*.

5.2. The syntactic category shifting approach

Dong (2005) analysis *hen* as the aspectual marker for adjectives in stand-alone sentences, just as *le* for verbs in independent sentences. According to him, in MC, *le* is a

perfective marker, while the reduplicative form of adjectives AABB-*de* and *shi...de* sequence are imperfective markers. Based on the fact that *hen* cannot co-occur with the above elements, as shown below, Dong (2005) draws the conclusion that *hen* is also an aspectual marker.

- (38) *Ta de lian hen hong le.
 she DE face very red LE
 ‘Her face turned very red.’
- (39) *Ta de lian hen hong-hong de.
 she DE face very red-red DE
 ‘Her face is red.’
- (40) *Ta de guandian shi hen cuowu de.
 She DE opinion SHI very wrong DE
 ‘Her opinion is wrong.’

As for why degree morphemes can be aspect marker, Dong’s explanation is that since degree morphemes involve comparing the degree of the state of the object indicated by the adjective with that of the standard of comparison (Kennedy 2007), the state of the object must be existent in the span of the time that includes the time of the comparison. However, this explanation is very stipulative.

Also, the author’s grammatical judgements of the sentences are problematic. Sentence (38) is fine in the context below, though there is a slight change in meaning:

- (41) Ta de lian hen hong le, bie zai rang ta hejiu le.
 She DE face very red LE don’t again let her drink LE
 ‘Her face is already very red, do not let her drink wine anymore.’

The following sentence is completely fine, too.

- (42) Ta de guandian shi hen zhengque de.
 She DE opinion SHI very correct DE
 ‘Her opinion is correct.’

Since *hen* can be a real degree morpheme (Li & Thompson 1981, Chui 2000, etc.), it is possible that sentences (38), (39) and (40) are bad due to semantic reasons. For instance, *honghong de* in (39) already denotes a high degree of redness, and there is no need for the degree intensifier *hen*.

When talking about why in contrastive situations, *hen* is not needed, Dong (2005) mentions that a contrast in itself is a comparison. Following his argument that comparison is made between two degrees, the adjectives need to be imperfective in order to be existent in the span of the time that includes the time of the comparison. As a matter of fact, this idea is consistent with my proposal that the appearance of *hen* creates a set of different degrees, with the cardinality of at least 2.

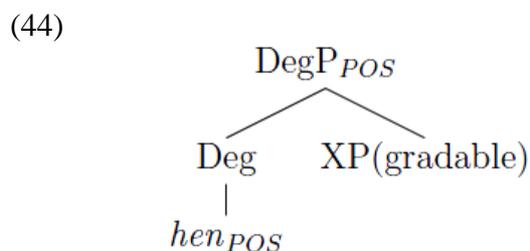
Grano (2011) proposes that degree adverbs, focus, etc., have the ability of turning adjectives into verbal categories. They are required to check the [+V] feature on T, assuming that T is always projected in clauses. According to him, the reason why *Zhangsan gao* is ungrammatical is that the adjective *gao* ‘tall’ fails to check the [+V] feature of T, by contrast, in *Zhangsan hen gao*, the use of the degree adverb *hen* ‘very’ licenses the [+V] feature, and thus enables the adjective *gao* to function as a predicate of T.

This analysis correctly captures the fact that when T is not projected, for example, in prenominal modification constructions and embedded clauses, *hen* is not obligatory. But it is not clear what the nature of the [+V] feature of T is. Also, it is not discussed how elements such as quantity phrases and *bi* phrases turn adjectives into verbal categories.

- (43) Zhangsan *bi* Lisi gao.
 Zhangsan BI Lisi tall
 ‘Zhangsan is taller than Lisi.’

The phrase *bi Lisi* in (43) is generally considered as a preposition phrase, it is not clear how it has the function of changing the categorial status of adjectives. Likewise, in the focus cases, it is even harder to understand how the phonological change can alter the syntactic categories of adjectives.

Syntactically, Zhang (2015) advances that *hen* is projected as the head of the functional projection DegP and s-selects a gradable phrase, either an AP or a stative VP. The structure is shown below:



She argues that *hen* is not a modifier or an adjunct but rather heads a DegP projection. Her key argument is that *hen* XP and XP have different syntactic behaviors. However, this argument is problematic. One of Zhang’s evidence is that in nominal exclusive constructions, once *hen* is added to the nominal XP, the sentence becomes unacceptable:

- (45) Wo ba (*hen) chengshi dangzuo yi zhong meide.
 I BA very honest regard-as one CL virtue
 ‘I regard being honest as a virtue.’

Zhang (2015) claims that this suggests that *chengshi* and *hen chengshi* are different syntactically, as the former can appear after *ba* but the other cannot. However, this argument does not make sense, because the reason why sentence (45) with *hen* is bad is that *chengshi*, in that case, is a noun, and degree words such as *hen* do not normally modify nouns. Sentence (45) should be glossed and translated in the following way instead:

- (46) Wo ba (*hen) chengshi dangzuo yi zhong meide.
 I BA very honesty regard-as one CL virtue
 ‘I regard honesty as a virtue.’

Another argument given by Zhang is that in nominal-exclusive positions, a bare VP is fine, but [*hen*+VP] is bad:

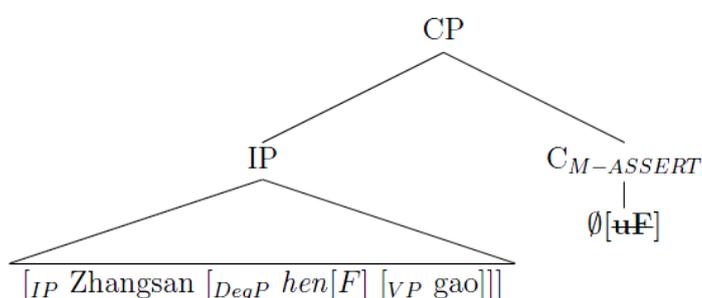
- (47) Wo ba (*hen) xihuan shige dangzuo yi zhong meide.
 I BA very like poem regard-as one CL virtue
 ‘I regard liking poems as a merit.’

However, as pointed out in Zhang (2015), nouns lack degree structures entirely, which means that NPs normally do not co-occur with degree words. This suggests that the reason sentence (47) is bad is not that the phrase *hen xihuan shige* cannot appear after *ba* but rather the phrase itself is illegal. It can be seen that Zhang’s two pieces of evidence for the arguments that *hen* XP and XP are syntactically different are problematic. Consequently, the proposal that *hen* heads a DegP projection does not hold.

5.3. The illocutionary force approach

Grano (2008) claims that degree morphology is needed only when the adjective is the entire predicate of the matrix-level declarative clause, to check the uninterpretable feature of $C_{M-ASSERT}$, which is the locus of the illocutionary force of the sentence. The schema is shown below.

(48)



According to Grano, the following sentence is fine because the clause *Zhangsan gao* is embedded: C is not projected, thus, no elements are required to check its feature.

(49) Wo zhidao [Zhangsan gao], dan mei xiangdao ta zheme gao.
 I know Zhangsan tall, but NEG expect he this tall
 ‘I knew Zhangsan was tall, but I didn’t expect he was this tall.’

However, as discussed in section 4.1, actually, sentence (49) is possible only because the two sub-clauses are connected by the contrastive coordinator *dan* ‘but’. Under the $\text{Pred}_{[+FOC]}$ analysis, it can be said that in (49), the [+FOC] feature of the Pred head in the embedded clause is checked by the coordinator *dan* ‘but’ (the contrast between ‘tall’ and ‘extremely tall’).

To summarize, in tackling the issue of the obligatory appearance of *hen* and other elements in predication constructions in MC, existing studies have limitations in two major aspects: (i) providing an analysis that covers adjectives both in prenominal modification position and postnominal predication position; (ii) unifying degree morphemes such as *hen*, focus intonation and the range of other elements that co-occur with the adjectives in predicate position. Both of these are captured by my proposed analysis.

6. Conclusion remarks and thinking points

The $\text{Pred}_{[+FOC]}$ analysis has important implications. Firstly, the contrast between (25) and (26a) above follows from this analysis. Secondly, it correctly predicts that when adjectives are used attributively or appear in small clauses, degree morphology is not needed; and conversely, when PredP is projected, adjectives should not be bare. Furthermore, this analysis coincides with the fact that non-gradable adjectives needs the accompany of the

shi...de sequence in predicate position, both suggesting that the claim that adjectives in MC are introduced as predicates by Pred_[+FOC] is on the right track.

To conclude, adjectives in MC are introduced as predicates by a functional projection PredP which bears an [+FOC] feature. Degree morphemes, focus intonation, *shi...de* sequence and other elements are required to check the [+FOC] feature by building contrastive scenarios. As for the difference between MC and English, it can be assumed that, in English, the copular *is* checks the [+FOC] feature of Pred, and accordingly, degree elements are optional. However, for future research, I plan to conduct a systematic investigation on contrastive constructions in MC in general and then show in detail how *hen* and other elements are linked to the focus interpretation in MC⁵.

References

- Adger, D. & Gillian, R. (2003). Predication and equation. *Linguistic Inquiry* 34, 325-359.
- Chui, K. (2000). Morphologization of the degree verb *hen*. *Language and Linguistics* 1, 45-59.
- Dong, H. (2005). Gradable adjectives in Chinese and degree morphemes. Ms. Cornell University.
- Grano, Thomas. (2008). Mandarin *hen* and the syntax of declarative clause typing. Ms. At LingBuzz.
- Grano, T. (2011). Mandarin *hen* and universal markedness in gradable adjectives. *Natural Language & Linguistic Theory* 30, 513-565.
- Huang, S. (2006). Property theory, adjectives, and modification in Chinese. *Journal of East Asian Linguistics* 15, 343-369.
- Kennedy, C. (1997). Antecedent-contained deletion and the syntax of quantification. *Linguistic Inquiry* 28, 662-688.
- Kennedy, C. (2007). Vagueness and grammar: The semantics of relative and absolute gradable adjectives. *Linguistics and Philosophy* 30, 1-45.
- Larson, K. (2009). Chinese as a reverse ezafe language. *Yuyanxue Luncong (Journal of Linguistics)* 39, 30-85.
- Lee, C. (2001). The Chinese negation marker *bu* and its association with focus. *Linguistics* 39, 703-731.
- Li, N. & Sandra, T. (1981). *Mandarin Chinese: A functional reference grammar*. Berkeley: University of California Press.
- Liu, L. (2010). The positive morpheme in Chinese and the adjectival structure. *Lingua* 120, 1010-1056.
- Paul, W. (2006). Zhu Dexi's two classes of adjectives revisited. In *Studies in Chinese language and culture*. In Anderl, C. and Eifring, H. (eds.), *Festschrift in honour of Christoph Harbsmeier on the occasion of his 60th birthday*, volume vii, 303-315. Oslo: Hermes Academic Publishing.
- Ramchand, G. (1996). Questions, polarity and alternative semantics. *Lingbuzz*.
- Rooth, E. (1992). A theory of focus interpretation. *Natural Language Semantics* 3, 75-116.
- Sproat, R. & Chilin S. (1991). The cross-linguistic distribution of adjective ordering restrictions. In Georgopoulos, C. and Ishihara R. (eds.), *Interdisciplinary approaches to language*, 565-593. Dordrecht: Kluwer.
- Sybesma, R. (1999). *The Mandarin VP*. Dordrecht: Kluwer.
- Tang, J. (1998). *Parametrization of features in syntax*. Doctoral dissertation, University of California.

⁵ I am grateful to one anonymous reviewer from TEAL-9 for pointing this out to me.

Zhang, N. (2015). Functional head properties of the degree word *hen* in Mandarin Chinese. *Lingua* 14-41.

*Fangfang Niu
Department of Linguistics (SLLF)
Queen Mary University of London
Mile End Road
London E1 4NS*

nfangfang@gmail.com

SYNTACTIC AND SEMANTIC CHANGE IN CHINESE

ALAIN PEYRAUBE
(CNRS, Paris, France)

Abstract

This paper discusses some of the most hotly debated topics over the past few years on syntactico-semantic change in a functional-cognitive perspective and propose a new model of grammatical change in Chinese by providing more solidly-based definitions of such notions as ‘grammaticalization’, ‘lexicalization’, ‘degrammaticalization’, ‘exaptation’, and ‘reanalysis’, as well as ‘analogy’ with respect to internal processes of change, but also for external ones, specifically, borrowing through language contact or contact-induced change.

It will propose that this model is constituted by just two internal mechanisms: reanalysis and analogy. Consequently, it will be argued that grammaticalization – which has been by far the focus of most of the studies on historical grammar in recent decades – is secondary. The processes of grammaticalization, lexicalization and exaptation will thus be viewed as sub-classes of reanalysis, while some ‘degrammaticalization’/lexicalization processes will be more aptly viewed as a sub-class of analogy.

The main motivations (if not genuine explanations) for grammatical change will also be discussed, i.e. semantic-pragmatic change, including mainly metaphorical extension, pragmatic inferencing or metonymization and (inter-)subjectification, as well as others, such as phonological change. Concerning the third – and external – mechanism of change, it will be shown that the several universals and principles of borrowing that have been proposed remain rather ill-defined.

1. Introduction *

The two essential mechanisms said to account for the appearance of new grammatical forms remain roughly those identified by Meillet (1912) a hundred years ago: analogy and grammaticalization.¹ Scholars working on diachronic syntax try to find answers to the four following questions: (i) ‘what motivates grammaticalization in the first place? (ii) what mechanisms lead to it? (iii) what are its probable paths of progression through time? (iv) what are its end results?’ (Hopper & Traugott 1993: 32). A third and external mechanism has to be necessarily added to these two internal ones: external borrowing, but it is principally the second of these two internal mechanisms, i.e. grammaticalization, that has been extensively discussed and commented on over the past thirty years.

* This research, presented during the 2nd Asian & European Linguistics Conference (AE-LINK2) at Newcastle University on 5-6 December 2014, has received funding from the French National Research Agency (ANR) for the project entitled ‘Typologie des processus synchronique et diachronique en Min-Sud (langue sinitique)’ (TYSOMIN n° 11-ISH2-001-01). It is based on a modified version of a paper presented at the Chinese University of Hong Kong CIEL-5 Conference (The 5th International Conference in Evolutionary Linguistics) held from 17-19 August, 2013, that will be published soon in the Proceedings of this Conference. I would like to thank Hilary M. Chappell, Cao Guangshun and Yu Hsiao-jung for their valuable comments.

¹ Meillet (1912) believed that there was a need to have a new class of changes that did not fit into any of the three categories of linguistic change allowed by the neo-grammarians and de Saussure (i.e. sound change, analogy, and borrowing) and so he was the first to coin the term ‘grammaticalization’.

The following issues have been much debated: (i) Does grammaticalization have any theoretical value?² (ii) Can the unidirectionality principle in grammaticalization be upheld and above all, is this unidirectionality of theoretical importance? (iii) Are there ‘degrammaticalization’ cases contradicting the unidirectionality principle, a term used widely today (Norde 2002, Heine 2003)? (iv) Is it necessary to apply the notion of exaptation, borrowed from biology, to linguistics, as suggested by Lass in 1990? (v) Are pragmatic inferencing (metonymization), metaphorical extension and subjectification the only main motivations for syntactic change, and hence the major mechanisms of semantic change?

In this paper, I would like to propose a coherent model of grammatical change in Chinese, by providing more solidly based definitions of such notions as grammaticalization, ‘degrammaticalization’, exaptation, but also of reanalysis and analogy, lexicalization, re-grammaticalization (Greenberg 1991), functional renewal, re-functionalization (Giacalone Ramat 1998), and hypo-analysis or under-analysis (Croft 2000), redefining their role and function in grammatical change.³

This model of grammatical change, roughly outlined in Peyraube (2005), has only two internal mechanisms of syntactic change: analogy and reanalysis. Grammaticalization, ‘degrammaticalization’ and exaptation are consequently secondary processes, and ‘degrammaticalization’ (or rather lexicalization as I will subsequently argue later that ‘degrammaticalization’ and lexicalization are one and the same phenomenon) and exaptation have to be distinguished: the first one belongs to both the mechanisms of analogy and reanalysis, while the last one belongs only to reanalysis. And there is still of course a third external mechanism: borrowing. Analogy (or generalization) thus comprises cases of ‘degrammaticalization’ / lexicalization, while reanalysis comprises grammaticalization, ‘degrammaticalization / lexicalisation and exaptation.

But what do we mean precisely by analogy, reanalysis, grammaticalization, unidirectionality and ‘degrammaticalization’ / lexicalization, and exaptation? What also could be the motivations – and not the genuine mechanisms – of the syntactic change?

2. Analogy (or Extension)

Hopper and Traugott (1993: 21) speak of ‘new paradigms (which) come into being through formal resemblance to already established paradigms.’ McMahan (1994: 71) defines analogical extension as follows: ‘generalization of a morpheme or relation which already exists in the language into new situations or forms.’ Another definition, which parallels the one usually given for reanalysis (see below), concerns the levels of structure. Analogy only modifies the surface structure and does not modify the underlying structure.

Analogy does not represent a principle of grammatical change: ‘the fact that many reanalyses can be interpreted as analogical extensions does not make analogy a principle of change, least of all an explanatory principle’ (Lightfoot, 1981: 225). As a matter of fact, analogy, as Kuriłowicz, (1945-9: 174), once said, is like rain water: it must take a certain path (channel, gutter, etc.) once it has fallen, but rain is not a necessity. In other words, analogy may tell us about the mechanism behind a change, but it cannot furnish the causes of such a change.

² Newmeyer (1998: 226, chapter ‘Deconstructing grammaticalization’) said: ‘there is no such thing as grammaticalization, at least in so far as it might be regarded as a distinct grammatical phenomenon requiring a distinct set of principles for explanation.’

³ Different definitions are currently being used for all these notions (linguists love taking old terms and giving them new definitions for new purposes); the situation can be characterized today as being quite confused, as researchers rarely mean the same things when they discuss what are apparently the same notions.

As for syntactic change in Chinese, almost everybody who has worked on Chinese historical syntax since the beginning of the 1980s has made use of analogical processes to account for grammatical change. This is not surprising, as almost all changes have an analogical ingredient, as stated by Anttila (1977), Lighfoot (1981), Kiparsky (1992) or Blevins and Blevins (2009). Chinese linguists, however, have often wrongly assumed that analogy (*lèitūī*) is more a factor motivating syntactic change than a simple mechanism of change.

Very recently, Kiparsky (2005, 2012), in a rethinking of the concept of analogy, proposed that analogical change is an optimization of grammar, the elimination of unmotivated grammatical complexity or idiosyncrasy. Understanding analogy as grammar optimization allows various types of analogy to be admitted: proportional, non-proportional, and even non-exemplar-based. He actually distinguishes two types of analogy: (i) exemplar-based analogy (comprising proportional analogical change and non-proportional analogical change); (ii) non-exemplar-based analogy, which seems a priori a contradiction in terms, as analogy by definition needs a model.

Examples of non-exemplar-based analogical changes are fusions of two words into one, which occur spontaneously without any particular model. It is driven by a language-independent preference for structural economy: other things being equal, one word is always better than two. The opposite process, fission of one word into two words is always exemplar-based, it occurs by analogy with specific existing constructions.

There are many examples of fusions or bonding (cases of two monosyllables having contracted into a single syllable written with one character) and fissions (reverse cases) in the history of Chinese that can be characterized as non-exemplar-based analogies (for the fusional processes) and non-exemplar-based analogies (for the fissional ones). Among the contractions of this kind in Classical Chinese (Late Archaic, (5th – 2nd centuries BCE), can be found:

- (1) *zhī* ‘third personal pronoun’ + *yú* ‘to, at’ > *zhū*; *bù* ‘not’ + *zhī* ‘third personal pronoun’ > *fú*; *wú* ‘not’ + *zhī* ‘third personal pronoun’ > *wù*; *yú* ‘at, to, from’ + *zhī* ‘third personal pronoun’ > *yān*; *hú* ‘why’ + *bù* ‘not’ > *hé*

Later, starting in the Pre-Medieval (Han) period, ca. 1st century BCE, all the above examples of fusion disappeared and a fission process came into operation, returning to the situation which held sway in the Late Archaic period.

- (2) *zhū* > *zhī* + *yú*; *fú* > *bù* + *zhī*; *wù* > *wú* + *zhī*; *yān* > *yú* + *zhī*; *hé* > *hú* + *bù*

Kiparsky goes even further, considering that grammaticalization is actually analogy, even though it is a special type of analogy, with the property that it is driven only by general principles and constraints of the language. It is a case of non-exemplar-based analogical change, allowing for new patterns to arise in language; while cases of ‘degrammaticalization’ (see further the definition) are better viewed as ordinary analogical changes of the exemplar-based type.

3. Reanalysis

Reanalysis is a new concept, compared with grammaticalization, but it also has already been used widely, at least since Langacker (1977) who gave the following definition: ‘Change in the structure of an expression or class of expressions that does not involve any immediate or intrinsic modification of its surface manifestation.’

Harris and Campbell (1995: 61), who were the first to consider reanalysis, instead of grammaticalization, as one of the two major internal mechanisms of syntactic change, adopted the definition of Langacker (1977). As this could also very well characterize grammaticalization, in opposition to analogy, which is a 'change in the surface manifestation of a syntactic pattern that does not involve immediate or intrinsic modification of underlying structure,' the two notions of grammaticalization and reanalysis have sometimes been confused, especially by the Chinese linguists. It is necessary to distinguish them. Even if most grammaticalization cases are also reanalyses, this is not always the case.

Consequently, I have adopted the following definition by Hagege (1993: 62): 'An operation by which language builders cease to analyze a given structure as they did previously, and introduce a new distribution of, and new relations between, the syntactic units that constitute this structure.' As a matter of fact, only Hagege's definition allows to consider major typological shifts such as word order change (OV > VO) as cases of reanalysis.

4. Grammaticalization

Meillet(1912)'s definition of grammaticalization ('*l'attribution du caractère grammatical à un mot jadis autonome*' - 'the shift of an independent word to the status of a grammatical element') has been improved, first by Kuryłowicz (1965): '(It) consists in the increase of the range of a morpheme advancing from a lexical to a grammatical or from a less grammatical to a more grammatical status.' Hopper and Traugott (1993: xv) define grammaticalisation as 'The process whereby lexical items and constructions come in certain linguistic contexts to serve grammatical functions, and, once grammaticalized, continue to develop new grammatical functions'. The 'process' becomes a single 'strong tendency' in the 2nd edition monograph (Hopper and Traugott, 2003: 231): 'Grammaticalization is a robust tendency for lexical items and constructions to be used in certain linguistic contexts to serve grammatical functions, and once grammaticalized, to be used to further develop new grammatical functions.'⁴

For Greenberg (1991: 303), however, grammaticalization is not only a shift from lexical to grammatical. It is '(a) process of development of grammatical elements from all sources'. Many Chinese linguists have preferred to take up this broad definition. However, I believe it is better not to see grammaticalization as an equivalent to grammatical change, and therefore it should not be considered as '(a) development of grammatical elements FROM ALL SOURCES'.

Actually, confined to diachronic studies as it should be understood, grammaticalization is a concept familiar to all Chinese linguistics specialists. At least from the Yuan dynasty onwards (1279-1368), Chinese scholars have observed that empty (grammatical) words were full (lexical) items in ancient times (see Sun 1996: 11, Hong B. 1998). Zhou Boqi, a Yuan dynasty scholar, wrote: *jīn zhī xū zì jiē gǔ zhī shí zì* 'today's empty words are all former full words.'

Special features of grammaticalization pathways, typical of Sinitic languages are given in Chappell and Peyraube (2011), with a discussion of the evolution of classifiers, and principally of the disposal or object marking constructions, passive and causative constructions, all having arisen out of serial verb constructions, creating new forms with complex predicates.

A lot of work has been done on grammaticalization since the 1990s and several principles (or tendencies) have been suggested, as the four typical 'heuristic principles'

⁴ See Traugott (2001). The continued development of already grammaticalized forms into new grammatical functions is what Givón 1991 called 'secondary grammaticalization.'

proposed by Hopper (1991): Layering, Divergence, Specialization, and Persistence.⁵ It is probably exaggerated to call these tendencies associated with unidirectionality 'heuristic principles,' – they are no longer called 'principles' in Hopper and Traugott (1993: 113) – but many good examples (in French, English, Malay, Hindi, etc.) have been found to illustrate them. They have also been effectively applied to Chinese syntactic change by Peyraube (1986, 1988, 1989) and Sun (1996: 165 *sq.*) for the history of the complementizer particle *de*, of the differential object marker *bǎ* and of the dative constructions. See also Heine and Kuteva (2002).

Another strong tendency in grammaticalization is that lexical meanings subject to grammaticalization are usually quite general. For example, verbs which grammaticalize, tend to be superordinate terms (hyperonyms) in lexical fields, for example *say*, *move*, *go*. They are not selected from specialized terms such as *whisper*, *assert*. etc. In other words, lexical items that grammaticalize are typically those known as 'basic words' or, at least, those which are easily accessible. Hagège (1993: 212) has even proposed a MGMTF 'principle' (the 'more general, more frequent principle'): 'due to the needs of human intercourse, Language Builders tend to use on a large scale the words that express general notions corresponding to everyday life.' This tendency has also been observed in Chinese grammaticalization: of hyperonymic verbs *yǔ* 'to give,' *bǎ* 'to take,' *zài* 'to be at,' *liǎo* 'finish' into the prepositions *yǔ* 'to,' *bǎ* 'pre-verbal object marker,' *zài* 'at,' *le* 'aspectual marker.'⁶

Basic to work on grammaticalization is also the concept of the cline. Forms do not shift abruptly from one category to another, but go through a series of gradual transitions, transitions which tend to be similar in type across languages. Most linguists will agree, for instance, that there is a 'cline of grammaticality' of the following type: content item > grammatical word > clitic > inflectional affix. Other examples have been provided in Chinese, for example V *liǎo* 'finish' > Phase complement > Aspectual suffix (see Mei 1994), or the evolution of *gong* from a lexical verb 'to share with' to an adverb 'together', and to a preposition 'with,' and to a conjunction 'and': V > Adverb > Preposition > Conjunction (see Liu and Peyraube 1994).

Another insight strongly developed by von der Gabelentz (1891: 251) is that syntactic change is not a linear process, but rather a cyclical one, or more exactly one that involves a spiral movement, in which changes do not exactly replicate themselves but parallel earlier changes in an approximate manner. Actually, there is little or no evidence to assume that the syntax of languages is developing in one direction through non-reproducible changes. The reason for the cyclical nature of change is to be found in the dialectical relationship between opposite needs of communication: ease of production and perception on the one hand, and the search for expressiveness on the other hand. A good and classic example of a cyclical change, confirmed again in Fleischman (1982) is the evolution of the future in Romance languages

⁵ Layering: 'Within a broad functional domain, new layers are continually emerging. As this happens, the old layers are not necessarily discarded, but may remain to coexist with and interact with the newer layers.'

Divergence: 'When a lexical form undergoes grammaticization (grammaticalization) to a clitic or affix, the original lexical form may remain as an autonomous element and undergo the same changes as ordinary lexical items.'

Specialization: 'Within a broad functional domain, at one stage a variety of forms with different semantic nuances may be possible; as grammaticization (grammaticalisation) takes place, this variety of formal choice narrows and the smaller number of forms selected assume more general grammatical meanings'.

Persistence: 'When a form undergoes grammaticization from a lexical to a grammatical function, so long as it is grammatically viable some traces of its original lexical meanings tend to adhere to it, and details of its lexical history may be reflected in constraints on its grammatical distribution.'

⁶ On the interactions between productivity and frequency, and on the differences between 'type frequency' and 'token frequency', see Brinton and Traugott (2005: 17).

(see Marchello-Nizia 1995: 29), where the analytic future-subjunctive of the Indo-European becomes a synthetic form in Classical Latin (*amabo, legam*), then analytic again in Vulgar Latin (*amare, habeo/volo*), and finally synthetic in form in Romance languages that have evolved from Latin, such as French (*aimerai*). Another example is the change Topic > Subject > Topic attested in several languages, including Chinese (Hagège 1978).

The idea of cyclic changes has been interestingly incorporated in various attempts at the matter by Cai (1986), Cao (1987a, 1987b, forthcoming), Mei (1984, 1987), D'Ésirat & Peyraube (1992). They noticed that some lexical items or grammatical morphemes attested at a certain time disappear during one or several centuries, before reappearing later. For instance, the adverbs *jiù* 'then' and *kuài* 'rapidly' were used during the Southern Song (12th and 13th centuries), to be later replaced by *bì'ān* and *jíu* under the Yuan (13th and 14th c.), and used again at the beginning of the Ming (15th c.). The demonstratives *zhè* 'this' and *nà* 'that' could be used alone as subjects under the Southern Song and again under the Ming (14th – 17th c.), but this was not the case during the Tang (7th – 10th c.) and the Yuan, where they had to be followed either by the classifier *gè* 个, or by the determinative particle *de* 的. The structure 'V + *le* 了 + O + *le* 了' (where the first *le* 了 is an aspectual marker and the second one a final particle) is attested under the Southern Song, after which it disappeared during the Yuan, and reappeared under the Ming.

But the most discussed 'principle' of grammaticalization, up to now, is certainly the 'unidirectionality principle' (see below section 6).

5. Exaptation

The term exaptation is widely used in work on evolution of language, but also now in historical morphosyntax. It was apparently first used in evolutionary biology by Gould & Vrba (1982): 'We wish to restrict the term *adaptation* only to those structures that evolved for their current utility; those useful structures that arose for other reasons, or for no conventional reasons at all, and were fortuitously available for other changes, we call *exaptations*' (Gould 1983, cited in Lass 1990: 80). Lass has been the first to suggest that the concept of exaptation could be used metaphorically to account for changes in grammar. He gave the following definition: 'Exaptation ... is the opportunistic co-optation of a feature whose origin is unrelated or only marginally related to its later use ... In other words (loosely) a conceptual novelty or invention' (1990: 80). Actually, when a form loses its function, or is only marginal within the system, only three possibilities arise: (i) it can be lost; (ii) it can be kept as marginal garbage; (iii) it can be reused for something else (that is exaptation).

Giacalone Ramat (1998) defines exaptation as 'refunctionalization under conditions of discontinuity in the developmental continuum'. Another term with reference to a somewhat similar phenomenon is 'regrammaticalization' (Greenberg 1991), but what is typically treated today as exaptation is the re-use of a marginal morpheme with an old function which becomes a more central morpheme with a new function (Traugott's 2004 formulation).

Exaptation, being a conceptual invention, is a special case of reanalysis. It has nothing to do with extension or analogy, or reformulation of paradigms in accordance with a target or a model. Greenberg (1991) or Giacalone Ramat (1998) argued that grammaticalization and exaptation are conflicting types of change. Norde (2002), Heine (2003), Traugott (2004) view them as essentially similar, but with different outcomes. In my view, it is neither grammaticalization nor 'degrammaticalization'. It is something different, as it is the reuse of an old form A for something else completely new (B), with no direct or indirect connection between A and B.

Many examples of exaptation have been raised and discussed for Indo-European languages. Lass (1990, 1997: 317-319, 2000) cited the following ones: the number/gender

agreement with nouns in Dutch adjective morphology has been reused as a marker on morphologically complex adjectives in Afrikaans; the Indo-European aspectual system has been reused for the Germanic tense system; the Old English present participle *-ende* has become the Modern English progressive form with *-ing*. Other cases are: the use of the Latin suffix *-ille* for the Romance definite article and clitic (e.g. French *le*); the re-emergence, with new functions, of the Old English conclusive perfect (*I have a letter written*) in the 17th century, to coexist with the perfect (*I have written a letter*) (Brinton & Stein 1995); the re-use twice of the Indo-European prefix *-sk* (originally for forming present tense): a first time as an inchoative in Latin (*pallesco* ‘grow pale’), and a second time as an affix in French (*je finis / nous finissons*) (Giacalone Ramat 1998: 111).

As for Chinese, very few cases have been reported until now, as the Chinese linguists have not paid much attention to the phenomenon of exaptation. One can nevertheless cite the reuse of the Ancient Chinese modal particle *yě* as an adverb in Medieval Chinese where *yě* means ‘also’. Another example could be the reuse of *nà*, which was a preposition meaning ‘at, to’ (equivalent to *yí*) in the Buddhist texts of the Late Han and Six dynasties period (2nd – 6th c. CE), as the distal demonstrative pronoun *nà* ‘that’, which began to appear under the Tang (7th – 10th c. CE).⁷

6. Unidirectionality and Lexicalization / Degrammaticalization

The ‘unidirectionality principle’ is claimed to be the main characteristic of grammaticalization (but not reanalysis), associated with it since the beginning. Haspelmath (1999) claimed: ‘Grammaticalization is irreversible’: the move is always from content words (or full words in Chinese *sh ě í*) to empty words (*xūcí*) or from lexical items to grammatical elements, or from less grammatical to more grammatical, never in the opposite direction. Given the theory of unidirectionality, it has been hypothesized that diachronically all minor categories (preposition, conjunction, auxiliary verb, pronoun, demonstrative, etc., i.e. relatively ‘closed’ categories) have their origins in major categories (relatively ‘open’ lexically, such as nouns and verbs).⁸ This assumption has been proved to be particularly valid in Chinese. It is however probably better to treat unidirectionality as a defining characteristic of grammaticalization (grammaticalization is unidirectional by definition) than a principle. See Lass (2000), Wu (2003).

Thus, the only direction of change involving a grammaticalization process is:

- (3) Discourse > Syntax > Morphology > Morphophonemics > Zero (Givón 1979: 209),
Lexical item > Grammatical element (Full word > Empty word), and Less
grammatical > More grammatical.

It is impossible to have: * Grammatical element > Lexical item, or * More grammatical > Less Grammatical.⁹

This being said, counterexamples to unidirectionality began to be discussed in the 1990’s, with well-documented instances of what has then been called ‘degrammaticalization’. Some have argued that such counterexamples are not damaging to the unidirectionality

⁷ Thanks to Cao Guangshun who has kindly suggested this potential case of exaptation.

⁸ See Talmy (2000: 22) for a definition of open classes (‘quite large and readily augmentable relative to other classes’) and closed classes (‘relatively small and fixed in membership’).

⁹ As grammaticalization is not an equivalent of grammatical change, but only one of its processes, claiming that grammatical items originate in lexical ones does not entail hypothesizing a language stage, at the origin, in which everything would be lexical, or that all languages will converge one day to have only grammatical elements.

hypothesis because they are sporadic (for example, Haspelmath 2004, Hopper and Traugott 2003, Wu 2003, 2005). They did not succeed, however, in trying to account for the exceptions to unidirectionality. Others (a growing number) have drawn the opposite conclusion and concluded that no special type of change such as grammaticalization even exists.

What is then in fact ‘degrammaticalization’? It is the opposite of grammaticalization, i.e. a change from a grammatical element to a lexical item (or from an empty word to a full word), or from a more grammatical element to a less grammatical one. It is a move from morphology to syntax, or from syntax to discourse. Some examples are:

- (4) English *up* [+ Preposition] > *up* [+ Verb] as in *up the ante*;
 English *-ism* [+ Suffix] > *ism* [+ Noun];
 English *dare* [+ Auxiliary verb] > *dare to* [+ Verb];
 English *calendar* [+ Noun] > *to calendar* [+ Verb];
 English *text* [+ Noun] > *to text* [+ Verb] as in *just text me*;
 French *trop* [+ Adverb] > *trop* [+ Adjective];
 French *pour* [+ Preposition] > *pour* [+ Noun];
 French *contre* [+ Preposition] > *contre* [+ Noun];
 Seto/Võru (South Estonian) *-ldaq* (+ Suffix) > *-ldaq* (+ Clitic) ‘without’;¹⁰
 Modern Greek *ksana-* [+ Prefix] > *ksana* [+ Clitic] ‘again’;
 Estonian *-p* [+ Prefix] > *ep* [+ Adverb];
 Spanish *-mos* [+ Suffix] > *nos* [first plural personal pronoun] (Janda 2001: 301);
 Irish *-muid* [+ Prefix] > *muid* [first plural personal pronoun].

In Chinese, where the grammaticalization process (within the reanalysis mechanism) is probably more important for grammatical change than in the Indo-European languages, the cases of ‘degrammaticalization’ are UNSURPRISINGLY rarer. One can nevertheless cite:

- (5) Chinese Adverb *tóng* ‘together’ > Noun *tóng* (such as *sān tóng* ‘the three together’);
 Chinese Suffix *huà* ‘-ization’ > Noun *huà* (as in *sì huà* ‘the four -izations, for the four modernizations’);
 Chinese Demonstrative pronoun *shì* ‘this’ > Copula, Verb *shì* ‘to be’;
 Chinese Demonstrative pronoun *zhī* ‘this’ > Verb *zhī* ‘to go’;
 Preposition *bǎ* (a pre-verbal object marker) > Verb *bǎ* for sentences like *wǒ bǎ nǐ zhìge hùchóng* [1SG – object marker – 2SG – this + CLASSIFIER – stupid prat] ‘You stupid prat!’ (see Jiang & Yang 2006).

The cases of fission raised above (in [2], section on Analogy) could also be considered as cases of ‘degrammaticalization’. As seen above, these cases of ‘degrammaticalization’ are subsumed by the mechanism of analogy, as ‘degrammaticalization’ cases are ordinary analogical changes of the exemplar-based type.

Actually, there is no need to talk of ‘degrammaticalization’ which violates the defining characteristic of unidirectionality for the grammaticalization process. All the cases of ‘degrammaticalization’ are *de facto* cases of lexicalization, which is another important process of language change. Van der Auwera (2002: 20), Ramat (1992, 2001) do not make any strict difference between ‘degrammaticalization’ and lexicalization.

¹⁰ Widely attested in Finno-Ugric, the promotion of the abessive case suffix (‘without’) to a clitic and even to an independent postposition has been triggered by analogy with its antonym, the comitative clitic *gaq* (‘with’).

Like grammaticalization, lexicalization, though far less systematically studied than grammaticalization, has been conceptualized in different ways.¹¹ In my view, it should not be defined as an ‘adoption of words into the lexicon’, as did Dong (2012), i.e. a process of word-formation, including, for Chinese, the main operations of compounding, derivation, reduplication. Being restricted, as suggested by Lehmann (1989, 2002) to ‘a process in which something becomes lexical’ (2002: 14) is not enough. The source of the process should be precisely indicated.¹² Lexicalization should be simply viewed as a historical process, as a reverse process of grammaticalization, as a development of concrete meanings from grammatical meanings, of full words (lexical items) from empty words (grammatical elements), a lexical item being a type of formal unit which belongs to the lexicon (that should not, in this case, be understood as an inventory of both ‘lexical’ and ‘grammatical’), sometimes called a lexeme, when typically contrasted with a grammatical morpheme or ‘gram’ (Brinton & Traugott 2005: 9-10)

7. External Borrowing

Analogy and reanalysis (including grammaticalization) are internal mechanisms of change. There is a third and external one, external borrowing through language contact. Borrowing, contrary to analogy, but like reanalysis, can introduce an entirely new structure into a language, and in this sense can produce a radical change. It is an ‘attempted reproduction in one language of patterns previously found in another’ (Haugen 1950 quoted in McMahon, 1994: 200).

Syntactic borrowing is probably both the least studied and the most abused area in syntactic change. Abused, because due to its explanatory power on the causes of change (its most obvious evidence being sheer necessity), this mechanism has often been evoked without a real identification of the source of the borrowing. Many examples can be cited for Chinese, as, for instance, the incorrect claim made by Song (1991) that the aspectual markers *le* or *zhe* have been borrowed from Altaic languages.

Little studied, because historical linguists have traditionally been strongly prejudiced in favor of internal mechanisms for linguistic changes. This is probably because the direction and extent of borrowing, and the kinds of features affected, are determined more by non-linguistic factors than by linguistic ones. The methodological inclination has thus been to consider the possibility of external causation only when all efforts to find an internal motivation have failed.

However, especially since the important works of Thomason and Kaufman (1988) and Thomason (2001), many scholars now consider that the possibility of borrowing as a cause for syntactic change should always be kept in mind. Li (1983, 1995) and Yue-Hashimoto (1993) have both, for example, emphasized borrowing as being an important force acting in favour of syntactic change in Chinese.

Several universals and principles of borrowing have been proposed, but they remain rather ill-defined. Constraints have also been identified (notably by Moravcsik 1978), but most of them appear to be too restrictive. Some of these debatable universals, principles and constraints are the following:

¹¹ For a good introduction to the lexicalization studies and a discussion of the different conceptualizations and definitions of lexicalization, see Brinton & Traugott (2005: 1-31), and Xing (2012: 1-19) and Dong X. (2012) for Chinese.

¹² The definition given by Anttila (1972: 151) ‘whenever a linguistic form falls outside the productive rules of grammar, it becomes lexicalized’ is not explicit about it, even if he considers that the ‘development of adverbs from nouns is also a case of lexicalization.’

Borrowing moves from the more to the less prestigious language. This condition is not absolute. We know many examples of Chinese structures borrowed from Altaic languages (Khitan, Jurchen, Mongolian, Mandchu) that have never had a more prestigious status than Chinese. Mei (1988) has thus convincingly showed that the opposition between *zánmen* 'we, inclusive' vs *wǒmen* 'we, exclusive,' which appeared in Chinese under the Jin (12th century), has been borrowed from Altaic languages, either Khitan or more probably Jurchen. Were these Altaic languages at that time more prestigious than Chinese? Probably not.

Basic vocabulary is only rarely affected. We nonetheless know that English borrowed a good deal of basic vocabulary from Norse: *sky*, *skin*, and even the pronouns *they*, *them*, *their*.

Structural compatibility is supposed to be required. Weinreich (1953: 25), after Jakobson, has stressed that a language 'accepts foreign elements only when they correspond to its tendencies of development.' However, any insistence that grammatical borrowing happens only in situations of shared structural similarities is simply wrong. Many examples involve grammatical borrowing from typologically divergent languages. See Li (1983, 1985, 1995), Mei (1988) for borrowing into Chinese from Altaic languages. Some of the Sinitic languages of Northwestern China (such as Linxia, Tangwang, Gangou) have even borrowed cases (accusative, genitive, dative, locative, instrumental, etc.) from Altaic languages (Mongolic or Turkic languages). See Peyraube (forthcoming).

Some categories (lexical elements) are said to rank highest in terms of borrowability, others lowest, if borrowable at all (grammatical forms). This claim is also debatable. An absolute ranking of this nature provides little real satisfaction.

Basic patterns are difficult to borrow. But a rather large number of cases have been reported in which basic word order patterns have in fact been borrowed. Faarlund (1990: 84) even claims that 'all known instances of a change from VO to OV are due to contact with OV languages.' Thomason and Kaufman (1988: 55) also believe that the word order of a language is one of the syntactic features that is very easily borrowable.

To sum up, if borrowing must indeed be accorded a more significant position among the three mechanisms of syntactic change, it is safer to consider all the proposed universals and principles of borrowing as general tendencies, instead of as absolute constraints. And it could be the case that language contact and the outcome of borrowing could have, at best, a trigger effect, releasing or accelerating grammatical phenomena which evolve independently.

Concerning the mechanism of external borrowing in grammatical change, the model of contact-induced grammatical change presented in Heine and Kuteva (2003, 2005) has begun to be applied in Sinitic and other East Asian and Southeast Asian languages by Wu (2013), Peyraube (forthcoming). The term 'borrowing' has been replaced by the term 'transfer'. There are transfers of grammatical meanings and structures from one language (the model language M) to another language (the replica language R) leading to grammatical replications. A contact-induced transfer is defined as follows: 'If there is a linguistic property X shared by two languages M and R, and these languages are immediate neighbors and/or are known to have been in contact with each other for an extended period of time, and X is also found in languages genetically related to M but not in languages genetically related to R, then we hypothesize that there is an instance of contact-induced transfer, more specifically that X has been transferred from M to R.' (Heine & Kuteva 2005: 33).

Over the past five years, a good amount of research on contact-induced grammatical change has been undertaken concerning Chinese and Altaic languages (Khitan, Jurchen, Mongolian and Manchu) that have been particularly important during the Liao (907-1125), Jin (1115-1234), Yuan (1206-1368), and Qing dynasties (1644-1911). See Cao & Chen (2009), Cao (2012), Djamouri (Forthcoming), Yu (2011, 2013), Yang (forthcoming a, forthcoming b), Zu (2007, 2009), Zhao (2010). A comparative analysis has also been made, on the largest attainable scale, of the grammatical transfers identified in historical documents

with those observable today in several Sinitic languages and Altaic languages (Turkic, Mongolic and Tungusic) in Northwestern China.

This research has thus also contributed to a better understanding of the typological characteristics that distinguish different Sinitic languages, as proposed by Chappell (2007) and of the issues involved in linguistic areas (see Chappell, forthcoming), reinforcing the links that exist between historical linguistics (and especially grammatical change) and typological linguistics.

8. Motivations of the grammatical change

Concerning the motivations of grammatical change, and no longer the mechanisms discussed above, there has not been much progress in research on this issue during the last decade. The main motivation has been said to be semantic-pragmatic change, itself refined by the three following mechanisms: metaphorical extension, pragmatic inferencing (or metonymization), and subjectification (and inter-subjectification).

As stated by Brinton and Traugott (2005: 28-29), subjectification is important as ‘the development of grammatical forms conceptually involves the recruitment of material to express the grammatical relations that the speaker envisions.’ As for metonymy, it is a cognitive process in which ‘one conceptual entity ... provides access to another conceptual entity.’ (Kövekşes & Radden 1998: 39). See Traugott and Dasher (2002), Peyraube (2005).¹³

If one looks carefully at the three mechanisms of grammatical change one by one, it can be assumed that analogical change has the two following main motivating factors: (i) A > B motivated by the abnormality of complexity of A, or by the generality or simplicity of B (pull/push model); (ii) semantic-pragmatic change, especially the metaphorical extensions that have been very often invoked in early work on grammaticalization when semantic change was considered (see Peyraube & Li 2008, 2012 for the semantic-pragmatic change in Chinese as a motivation of the mechanisms of analogy). The motivations for reanalysis are also (i) semantic-pragmatic change, but more pragmatic inferencing or metonymisation than metaphorical extension, which is more related to analogy; (ii) subjectification and/or inter-subjectification (Traugott 2004), (iii) other motivations such as phonological change, typological tension and or structural requirements apply to both analogy and reanalysis, but they are secondary. They did not play an important role in Chinese grammatical change and some of them (especially structural requirement and typological tension) should probably be better viewed as mechanisms and not motivations. Finally, the main motivation for external borrowing is obviously language contact.

9. Conclusion

There are two – and only two – powerful internal mechanisms of grammatical change proposed in this paper: analogy and reanalysis. In this model, grammaticalization, ‘degrammaticalization’, and exaptation are secondary operations, included either in analogy or in reanalysis. There is a third important mechanism, but one that is external: borrowing. Analogy (or extension, or generalization) comprises what is often called ‘degrammaticalization’ (see the fission examples above) that are actually lexicalization. Reanalysis comprises grammaticalization, exaptation, but also most of the lexicalization cases

¹³ Liu, Cao & Wu (1995) identified four motivating factors of grammatical change in Chinese: change in syntactic position, semantic change, pragmatic influence, reanalysis. It is not clear, however, why they considered the change of syntactic position and above all reanalysis as a motivation of grammaticalization. See also Cao (2001), Jiang (2004), Mei (1994, 1996), Wu (2005).

that belong to grammatical change (since the sources of the change are grammatical morphemes), as well as lexical change.

Analogical change has multiple motivating factors: (i) A > B can be motivated by the abnormality or complexity of A, or by the generality or simplicity of B (pull/push model); (ii) the main motivation is semantic-pragmatic change, especially on the form of metaphorical extensions. Structural ambiguity by itself does not trigger analogy.

The main motivation for reanalysis is also semantic-pragmatic change, but more in terms of pragmatic inferencing (or metonymisation) than in terms of metaphorical extension (which is more related to analogy). Another mechanism of semantic-pragmatic change, that plays an important role as a motivation of reanalysis is subjectification (and inter-subjectification).

The main motivation for external borrowing is obviously language contact. Finally, we should be more aware of the contact-induced change, as the situation today is about the same as it was ten years ago, when Traugott (2001) wrote: ‘Most of analyses up to now have been conducted largely in the context of putative homogeneous developments. When we look at contact situations, complications arise.’

References

- Akimoto, M. (2004). *Linguistic Studies based on corpora*. Tokyo: Hituzi Syobo Publishing C^o.
- Andersen H. (ed.) (1995). *Historical Linguistics 1993: Selected Papers from the 11th ICHL, Los Angeles 16-20 August 1993*. Amsterdam/Philadelphia: John Benjamins.
- Anttila, R. (1972). *An Introduction to Historical and Comparative Linguistics*. London: MacMillan Publishing C^o.
- Anttila, R. (1977). *Analogy*. The Hague: Mouton.
- Auwera, J. van der (2002). More thoughts on degrammaticalization. In Wisner & Diewald (eds.), 19-29.
- Blevins, J. & Blevins J. (2009). *Analogy in Grammar*. Oxford: Oxford University Press.
- Brinton, L. & Stein D. (1995). Functional renewal. In Andersen (ed.)
- Brinton, L. & Traugott, E.C. (2005). *Lexicalization and Language Change*. Cambridge: Cambridge University Press.
- Bright, W. (ed.), *International Encyclopedia of Linguistics, vol. I*. New York: Oxford University Press.
- Burridge, K. & Blake, B. (eds.) (2003). *Historical linguistics 2001 - Papers from ICHL 15*. Melbourne, Australia.
- Cai, J. (1986). Shi bei (On bei). *Zhongguo yuwen* 4, 314-315. [in Chinese]
- Cao, G. (1987a). Shi shuo jiu he kuai zai Song dai de shiyong ji youguan de duandai wenti (On the use of jiu and kuai under the Song and some issues of dating). *Zhongguo yuwen* 4, 288-294. [in Chinese]
- Cao, G. (1987b). Yuqici le yuanliu qianshuo (On the origin of the modal particle le). *Yuwen yanjiu* 2, 10-15. [in Chinese]
- Cao, G. (2001). Hanyu yufa shi yanjiu zhong de lishi shijiao (An historical visual angle in research on Chinese historical grammar). MS [in Chinese]
- Cao, G. (2012). Xibei fangyan teshu yufa xianxiang yu Hanyu shi zhong yuyan jiechu yinfa de yufa gaibian – yige fanchou weili (Special syntactic features in Northwestern dialects and contact-induced syntactic change in the history of Chinese – for example the case category). *Lishi yuyanxue yanjiu* 5, 85-94. [in Chinese]
- Cao, G. (Forthcoming). Jindai hanyu zhong de nazuomo (On the use of nazuomo in Modern Chinese). *Zhongguo yuwen*. [in Chinese]

- Cao, G. & Chen, D. (2009). Yuan baihua teshu yuyan xianxiang zai yanjiu (A study of the special syntactic features in Yuan vernacular). *Lishi yuyanxue yanjiu* 2, 108-123. [in Chinese]
- Cao G., Chappell H., Djamouri D. & Wiebusch T. (eds.) (2013). *Breaking down the barriers: Interdisciplinary studies in Chinese linguistics and beyond*. Taipei: Academia Sinica.
- Cao, G., Djamouri, R. & Peyraube, A. (eds.) (Forthcoming). *Languages in contact in Northwestern China*. Paris: Collection des Cahiers de Linguistique Asie Orientale.
- Chappell, H. (2007). Hanyu fangyan de chuzhishi biaoji de leixing (A typology of object-marking construction: a pan-Sinitic view. *Yuyanxue luncong* 36, 184-209. [in Chinese]
- Chappell, H. (Forthcoming). Languages of China in their East Asian context', in Hickey R. (ed).
- Chappell, H. & Peyraube, A. (2011). Grammaticalization in Sinitic Languages. In Narrog & Heine (eds.), 786-796.
- Croft, W. (2000). *Explaining language change: an evolutionary approach*. Harlow, Essex, UK: Pearson Education.
- Désirat, M. & Peyraube, A. (1992). Some remarks on the demonstratives in the Fuzhou dialect with reference to their historical evolution in Medieval and Modern Chinese. *Chinese languages and linguistics* 1, 493-522.
- Djamouri, R. (Forthcoming). Reanalysis of the modal adverbs *shuo* and *shuozhe* in Tangwang. In Cao, Djamouri & Peyraube (eds.)
- Dong, X. (2012). Lexicalization in the history of the Chinese language. In Xing (ed.), 235-274.
- Faarlund, J. (1990). *Syntactic change: toward a theory of historical syntax*. Berlin: Mouton de Gruyter.
- Fischer, O., Rosenbach, A. & Stein, D. (2000). Pathways of change – grammaticalization in English. Amsterdam: John Benjamins.
- Fischer, O., Norde, M. & Perridon, H. (eds.) (2004). *Up and Down the Cline: the Nature of grammaticalization*. Amsterdam/Philadelphia: John Benjamins.
- Fisiak, J. (ed.) (1985). *Historical Semantics, Historical Word-Formation*. Berlin: Mouton de Gruyter.
- Fisiak, J. (ed.) (1995). *Linguistic Change Under Contact Conditions*. Amsterdam: J. Benjamins.
- Fleischman, S. (1982). *The future in thought and language. Diachronic evidence from Romance*. Cambridge: Cambridge University Press.
- Gabelentz, G. von der (1891). *Die Sprachwissenschaft: ihre Aufgaben, Methoden und bisherigen Ergebnisse*. Leipzig: Weigel.
- Giacalone Ramat, A. (1998). Testing the boundaries of grammaticalization. In Giacalone Ramat & Hopper (eds.), 107-127.
- Giacalone Ramat, A. & Hopper, P. (eds.), *The limits of grammaticalization*. Amsterdam/Philadelphia: John Benjamins.
- Givón, T. (1979). *On Understanding Grammar*. New York: Academic Press.
- Givón, T. (1991). The evolution of dependent clause morpho-syntax in Biblical Hebrew. In Traugott & Heine (eds.), 257-310.
- Gould, S.J. & Vbra E.S. (1982). Exaptation – a missing term in the science of form. *Paleobiology* 8, 4-15.
- Gould, S.J. (1983). *Hen's Teeth and Horse's Toe. Further Reflections in Natural History*. New York: Norton.
- Greenberg, J.H. (1991). The last stages of grammatical elements: contractive and expansive desemanticization. In Traugott & Heine (eds.), 301-314.

- Greenberg, J.H., Ferguson, C. & Moravcsik, E. (ed.) (1978). *Universals of Human Language*. Stanford: Stanford University Press.
- Hagège, C. (1978). Du thème au thème en passant par le sujet : vers une théorie cyclique. *La linguistique* 14, 3-38.
- Hagège, C. (1993). *The Language Builder*. Amsterdam/Philadelphia: John Benjamins.
- Harris, A. & Campbell, L. (1995). *Historical syntax in cross-linguistic perspective*. Cambridge: Cambridge University Press.
- Haspelmath, M. (1999). Why is grammaticalization irreversible? *Linguistics* 37, 1043-1068.
- Haspelmath, M. (2004). On directionality in language change with particular reference to grammaticalization. In Fischer, Norde & Perridon (eds.), 17-44.
- Haugen, E. (1950). The Analysis of Linguistic Borrowing. *Language* 57, 183-91.
- Heine, B. (2003). (De)grammaticalization. In Burridge & Blake (eds.)
- Heine, B. & Kuteva, T. (2002). *Word Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.
- Heine, B. & Kuteva, T. (2003). On contact-induced grammaticalization. *Studies in Language* 27-3, 529-572.
- Heine, B. & Kuteva, T. (2005). *Language Contact and Grammatical Change*. Cambridge: Cambridge University Press.
- Hickey, R. (ed). *The Cambridge Handbook of Areal Linguistics*. Cambridge: Cambridge University Press.
- Ho D., Cheung, S., Pan, W. & Wu, F. (eds.) (2006). *Linguistic Studies in Chinese and Neighboring Languages: Festschrift in Honor of Professor Pang-Hsin Ting on His 70th Birthday*. Taiwan: Academia Sinica.
- Hong, B. (1998). Lun Hanyu shici xuhua de jizhi (On the mechanisms of full words becoming empty words in Chinese). In Hong Bo, *Hanyu lishi yufa yanjiu (Research on Chinese Historical Grammar)*. Beijing: Shangwu yinshuguan, 299-310. [in Chinese]
- Hopper, P. (1991). On some principles of grammaticization. In Traugott & Heine (eds.), 17-36.
- Hopper, P. & Traugott, E.C. (1993). *Grammaticalization*. Cambridge: Cambridge University Press. 2nd edition in 2003.
- Hornstein, N. & Lightfoot, D. (eds.) (1981). *Explanations in Linguistics. The Logical Problem of Language Acquisition*. London: Longman.
- Janda, R. (2001). Beyond Pathways and Unidirectionality: on the discontinuity of language transmission and the counterability grammaticalization. *Language sciences* 22, 265-340.
- Jiang, L. & Yang, Y. (2006). Jushi shengsuo yu xiangguang de niyufahua qingxiang (Sentence ellipsis and tendencies to reverse grammaticalization. In Ho, Cheung, Pan & Wu (eds), 439-466. [in Chinese]
- Jiang, S. (2004). Hanyu yufa yanbian ruogan wenti de sikao (Reflections on main issues concerning the Chinese grammatical change). Paper delivered at the 5th International Conference on Ancient Chinese Grammar.
- Jonas, D., Whitman, J. & Garrett, A. (eds.) (2012). *Grammatical Change: Origins, Nature, Outcomes*. Oxford: Oxford University Press.
- Kiparsky, P. (1992). Analogy. In Bright (ed.), 56-61.
- Kiparsky, P. (2005). Grammaticalization as Optimization. Paper delivered at Santigao de Compostela. July 17.
- Kiparsky, P. (2012). Grammaticalization as Optimization. In Jonas, Whitman, & Garrett (eds.), 15-51.

- Kövecses, Z. & Radden, G. (1998). Metonymy: Development of cognitive linguistic view. *Cognitive Linguistics* 9, 37-77.
- Kuryłowicz, J. (1945-9). La nature des procédés analogiques. *Acta Linguistica* 5, 15-37.
- Kuryłowicz, J. (1965). The evolution of grammatical categories. Reprinted in Kuryłowicz, J. (1975). *Esquisses linguistiques* 2: 38-54. München: Fink.
- Langacker, R. (1977). Syntactic Renalysis. In Li (ed.), 57-139.
- Lass, R. (1990). How to do things with junk: exaptation in language evolution. *Journal of Linguistics* 26, 79-102.
- Lass, R. (1997). *Historical linguistics and language change*. Cambridge: Cambridge University Press.
- Lass, R. (2000). Remarks on (uni)directionality. In Fischer, Rosenbach & Stein (eds.), 207-227.
- Lehmann, C. (1989). Grammatikalisierung und Lexikalisierung. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 42, 11-19.
- Lehmann, C. (2002). New reflections on grammaticalization and lexicalization. In Wischer & Diewald (eds.), 1-18.
- Li, C.N. (ed.) (1977). *Mechanisms of syntactic change*. Austin: University of Texas Press.
- Li, C.N. (1983). Languages in contact in Western China. *Papers in East Asian Languages* 1, 31-51.
- Li, C.N. (1985). Contact-induced semantic change and innovation. In Fisiak (ed.), 325-337.
- Li C.N. (1995). Language contact in China: Is Mandarin Chinese derived from a pidgin? In Fisiak (ed.), 131-148.
- Lightfoot, D. (1981). Explaining syntactic change. In Hornstein & Lightfoot (eds.), 209-240.
- Liu, J. & Peyraube, A. (1994). History of some coordinative conjunctions in Chinese. *Journal of Chinese Linguistics* 22-2, 179-201.
- Liu, J. Cao G. & Wu F. (1995). Lun xiufa hanyu cihui yufahua de ruogan yinsu (On the main parameters triggering the grammaticalization of the Chinese lexicon). *Zhongguo yuwen* 3, 161-169. [in Chinese]
- Marchello-Nizia, C. (1995). *L'évolution du français. Ordre des mots, démonstratifs, accent tonique*. Paris: Armand Colin.
- McMahon, A. (1994). *Understanding Language Change*. Cambridge: Cambridge University Press.
- Mei, T. (1984). Cong yuyan shi kan ji ben Yuan zaju binbai de xiezuo shiqi (On the dating of the vernacular sections of the Yuan operas from a linguistic point of view). *Yuyanxue luncong* 13, 111-153. [in Chinese]
- Mei, T. (1987). Tang, Wu dai zhe, na bu dan yong zuo zhuyi (On zhe and na that cannot be used alone as subjects under the Tang and Five Dynasties). *Zhongguo yuwen* 3, 205-207. [in Chinese]
- Mei, T. (1988). Beifang fangyan zhong diyi rencheng daici fushu baokuoshi he paichushi duili de lai yuan (Origin of the opposition between inclusive and exclusive forms of the 1st plural personal pronouns in Northern Chinese). *Yuyanxue luncong* 15, 141-145. [in Chinese]
- Mei, Tsu-lin (1994). Tang dai Song dai gongtongyu de yufa he xiandai fangyan yufa (The grammar of the Tang and Song koine and the grammar of Modern Chinese dialects). *Chinese Languages and Linguistics* 2, 61-97. [in Chinese]
- Mei, T. (1996). Hanyu yufa shi zhong jige fanfu chuxian de yanbian fangshi (On several paths of change that occur repeatedly in Chinese grammatical change). Paper delivered at the 2nd International Conference on Ancient Chinese Grammar. Beijing University. [in Chinese]

- Meillet, A. (1912). L'évolution des formes grammaticales. In Antoine Meillet, *Linguistique historique et linguistique générale*. Paris: Champion, 1948, 130-148.
- Moravcsik, E. (1978). Language contact. In Greenberg, Ferguson & Moravcsik (ed.), 93-123.
- Narrog, H. & Heine, B. (eds.) (2011). *The Oxford Handbook of Grammaticalization*. Oxford: Oxford University Press.
- Newmeyer, F. (1998). *Language form and function*. Cambridge, MA: MIT Press.
- Norde, M. (2002). The final stages of grammaticalization: affixhood and beyond. In Wischer & Diewald (eds.), 45-66.
- Peyraube, A. (1986). Shuang binyu jiegou - cong Han dai zhi Tang dai de lishi fazhan (Double-object constructions – Historical development from Han times to the Tang dynasty). *Zhongguo yuwen* 3, 204-216. [in Chinese]
- Peyraube, A. (1988). Syntactic change in Chinese: on grammaticalization. *The Bulletin of the Institute of History and Philology* 59, 617-52.
- Peyraube, A. (1989). Zaoqi ba ziju de jige wenti (Some issues on the earliest disposal constructions). *Yuwen yanjiu* 1, 1-9. [in Chinese]
- Peyraube, A. (2005). Leitui, yufahua, qu yufahua yu gongneng gengxin (Analogy, Grammaticalization, Degrammaticalization and Exaptation). Paper presented at the third International Conference on Grammaticalization, Luoyang, China. October 27-28. [in Chinese]
- Peyraube, A. (Forthcoming). A comparative analysis of the case system in Northwestern Sinitic Languages. In Cao, Djamouri & Peyraube (eds.).
- Peyraube, A. & Li, Ming (2008). Yuyi yanbian lilun yu yuyi yanbian he jufa yanbian (Theoretical considerations on semantic change and the relations between semantic change and syntactic change). In Shen & Feng (eds.), 1-25. [in Chinese]
- Peyraube, A. & Li, M. (2012). The semantic historical development of modal verbs of volition in Chinese. In Xing (ed.), 149-168.
- Ramat, P. (1992). Thoughts on degrammaticalization. *Linguistics* 30, 549-560.
- Ramat, P. (2001). Degrammaticalization or transcategorization? In Schaner-Wolles, Rennison & Neubarth (eds.), 393-401.
- Schaner-Wolles, C., Rennison, J. & Neubarth, F. (eds.) (2001). *Naturally! Linguistic Studies in Honour of Wolfgang Ulrich Dressler Presented on the Occasion of his 60th Birthday*. Torino: Rosenbach and Sellier.
- Shen, Y. & Feng, S. (eds.) (2008). *Dangdai yuyanxue lilun he hanyu yanjiu (Contemporary Theoretical Linguistics and Research on Chinese Language)*. Beijing: Shangwu yinshuguan 2008.
- Song, J. (1991). Hanyu zhuci *le*, *zhe* yu Aerdai zhu yuyan de guanxi (The particles *le* and *zhe* and their links to Altaic languages). *Minzu yuwen* 6, 59-67. [in Chinese]
- Sun, C. (1996). *Word Order Change and Grammaticalization in the History of Chinese*. Stanford: Stanford University Press.
- Talmy, L. (2000). *Toward a Cognitive Semantics*. Cambridge, MA: MIT Press.
- Thomason, S. (2001). *Language contact*. Edinburgh: Edinburgh University Press.
- Thomason, S. & Kaufman, T. (1988). *Language Contact, Creolization and Genetic Linguistics*. Berkeley and Los Angeles: University of California Press.
- Traugott, E.C. (2001). Legitimate counter-examples to unidirectionality. Paper presented at Freiburg University, October 17th.
- Traugott, E.C. (2004). 'Exaptation and grammaticalization', in Akimoto (ed.), 133-156.
- Traugott, E.C. & Heine, B. (eds.) (1991). *Approaches to Grammaticalization*. Amsterdam: John Benjamins.
- Traugott, E.C. & Dasher, R. (2002). *Regularity in Semantic Change*. Cambridge: Cambridge University Press.

- Weinreich, Uriel (1953). *Languages in contact: findings and problems*. The Hague: Mouton.
- Wisher, I & Diwald, G. (eds.) (2002). *New Reflections on Grammaticalization*. Amsterdam/Philadelphia: John Benjamins.
- Wu, F. (2003). Guanyu yufahua de danxiangxing wenti (On the issue of unidirectionality in grammaticalization). *Dangdai yuyanxue* 4. Reprint in Wu Fuxiang, *Yufahua yu hanyu lishi yufa yanjiu* (*Grammaticalization and Researches on Chinese Diachronic Syntax*. Hefei: Anhui jiaoyu chubanshe. 24-49. [in Chinese]
- Wu, F. (2005). Hanyu yufahua yanbian de jige leixinxe tezheng (On some typological characteristics of the grammatical change in Chinese). *Zhongguo yuwen* 6, 483-494. [in Chinese]
- Wu, F. (2013). Yufa fuzhi yu jiegou yanbian (Grammatical replication and structural changes). In Wu & Xing (eds.), 329-359.
- Wu, F. & Xing X. (eds.) (2013), *Yufahua yu yufa yanjiu* (*Grammaticalization and research on Grammar*) 6. Beijing: Shangwu yinshuguan.
- Xing, J. Z. (ed.) (2012). *Newest Trends in the Study of Grammaticalization and Lexicalization in Chinese*. Berlin: de Gruyter Mouton.
- Yang, Y. (Forthcoming a). Liao-Song shiqi Hanyu yu Qidan yu de jiechu shuolüe (Language contact between the Chinese and Khitan languages during the Song and Liao dynasties)'. In Cao, Djamouri & Peyraube (eds.).
- Yang, Y. (Forthcoming b). Song-Jin shiqi Hanyu yu Nüzhen yu de jiechu shuolüe (Language contact between the Chinese and Jurchen languages during the Song and Jin dynasties). In Cao, Djamouri & Peyraube (eds.).
- Yu, H. (2011). Shiyong duibi fenxi *Lao Qida* zhong de teshu yuyan xianxiang (Contrastive analysis and the special features in *Lao Qida*. *Lishi yuyanxue yanjiu* 4, 1-17. [in Chinese]
- Yu H. (2013). Contact and change in the history of Chinese language. In Cao, Chappell, Djamouri & Wiebusch (eds.), 485-501.
- Yue-Hashimoto, A. (1993). The Lexicon in Syntactic Change: Lexical Diffusion in Chinese Syntax. *Journal of Chinese Linguistics* 21-2, 213-254.
- Zhao, C. (2010). Yuan-Ming shiqi *houtou* biao shijian de yongfa ji qi lai yuan (On the origin and function of *houtou* expressing time in Yuan-Ming Chinese). Paper delivered at the 4th International Symposium on Language contact and Northwestern Chinese dialects. Santa Barbara, University of California at Santa Barbara, November 29 – December 1st. [in Chinese]
- Zu, S. (2007). Yuandai de Mengshi Hanyu jiqi shiti fanchou de biaoda – yi zhiyiti wenxian yanjiu wei zhongxin (On the Mongolian pidgin Chinese of Yuan dynasty with special reference to its tense and aspect category). *Bijiao yuyanxue* 1, 1-13. [in Chinese]
- Zu, S. (2009). Yuandai de *Han'er* yanyu (On the *Han'er* language of the Yuan dynasty). *Lishi yuyanxue yanjiu* 2, 124-135. [in Chinese]

Alain Peyraube
Centre National de la Recherche Scientifique, Paris, France
38 rue Bezout
75014 Paris
France

alain.peyraube@gmail.com

FACTS: THE INTERPLAY BETWEEN THE MATRIX PREDICATE AND ITS CLAUSAL COMPLEMENT *

JI YOUNG SHIM & TABEA IHSANE
(University of Geneva)

Abstract

The present study provides a minimalist account of diverse semantic and syntactic patterns found in three different types of clausal complements selected by factive and non-factive predicates in Korean. It will be shown that Korean challenges the widely-held dichotomous views on factive and non-factive sentential complements, such as the presence vs. absence of presupposition and the presence vs. absence of islandhood. We argue that the semantic and syntactic behavior (factivity and islandhood) of the clausal complement is determined by the interplay between two related factors, the type of the matrix predicate (factive vs. non-factive) and the type of the clausal complement in Korean.

1. Introduction

Since the seminal work by Kiparsky & Kiparsky (1970), FACTIVITY has been one of the extensively discussed topics in the fields of syntax and semantics (e.g., Hooper & Thomson 1973, Cattell 1978, Hegarty 1992, de Cuba & Ürögdi 2001, to name a few). Studies on factive predicates (e.g., *regret*, *know*, *remember*) vs. non-factive predicates (e.g., *believe*, *say*, *think*) have shown that the two types of predicates are distinguished from each other in several ways, including the presence vs. absence of the presupposition. As originally noted by Kiparsky & Kiparsky, a factive predicate presupposes the truth of its clausal complement, whereas a non-factive predicate does not, as shown in (1).

- (1) a. #Kibo regrets that Dana read this book, but in fact she didn't read it. *factive*
b. Kibo believes that Dana read this book, but in fact she didn't read it. *non-factive*

In (1a), the complement of the factive predicate *regret* is required to be true in order for the entire sentence to be felicitous. Thus, when the truth of the clausal complement is negated, the whole sentence is no longer felicitous. In contrast, the truth condition of the entire sentence is independent of that of the embedded clause with a non-factive predicate such as *believe* (Melvold 1991). Thus, when the truth of the clausal complement is negated, the whole sentence remains felicitous in (1b).

In addition, the sentential complements of factive and non-factive predicates behave differently with respect to syntactic movement. While subject, object, and adjunct extraction are all possible out of non-factive sentential complements (2), only object extraction is marginally allowed in factives in English (3) (Hegarty 1990, Varlakosta 1994, Basse 2008).

- (2) a. What does Kibo believe that Dana read ___?
b. Who does Kibo believe (*that) ___ read this book?
c. Why does Kibo believe that Dana read this book ___?

* This study is supported by a Swiss National Science Foundation grant (#100012_146699/1), which is gratefully acknowledged. We would also like to thank Marcel den Dikken, Michal Starke, and the audience at AE-Link 2 to discuss the content of the paper at various stages of its development.

- (3) a. [?]What does Kibo regret that Dana read ___?
 b. *Who does Kibo regret that ___ read this book?
 c. *Why does Kibo regret that Dana read this book ___?

While the distinction between factives and non-factives seems to be clear in English, Korean complicates the picture. Korean has THREE different types of sentential complements, (a) the *ko* clause, (b) the *kes* clause with an overt tense morpheme and a (declarative) force marker, and (c) the *kes* clause without an overt tense morpheme and a force marker. Examples are provided in (4).¹

- (4) a. Kibo-nun [Dana-ka i chayk-ul ilk-ess-ta-ko]
 Kibo-TOP Dana-NOM this book-ACC read-PAST-DECL-COMP
- b. Kibo-nun [Dana-ka i chayk-ul ilk-ess-ta-nun kes-ul]
 Kibo-TOP Dana-NOM this book-ACC read-PAST-DECL-LNK thing-ACC
- c. Kibo-nun [Dana-ka i chayk-ul ilk-un kes-ul]
 Kibo-TOP Dana-NOM this book-ACC read-LNK thing-ACC
 ‘But in fact, Dana did not read this book.’
- yukamsulewehay-ss-ta / mit-ess-ta
 regret-PAST-DECL believe-PAST-DECL
 ‘Kibo regretted/believed that Dana read this book.’

Korean is a head-final language (SOV order), and agglutinative in its morphology. As we see in (4a), the (bracketed) clausal complement is headed by the complementizer *ko*, in which the verb *ilk* ‘read’ is suffixed with the past tense morpheme *ess* and the declarative force marker *ta*. In (4b) and (4c), on the other hand, the embedded clause is headed by *kes*, which roughly translates into ‘(a) thing’ in English. However, the two *kes* clauses can be distinguished from each other; in (4b) the embedded verb *ilk* is fully marked with the past tense morpheme and the declarative force marker, similar to the verb *ilk* in the *ko* clause in (4a). However, the *kes* clause in (4c) is BARE in the sense that the verb *ilk* is in its bare form without being suffixed with an overt tense morpheme or a force marker. These three types of clausal complements can be embedded under both a factive predicate such as *regret* and a non-factive predicate such as *believe*.

Further analyses of the three sentential complements in (4) reveal that the presupposition of the complement is derived from the interaction of the matrix predicate and the type of complement, out of which intricate extraction patterns emerge, which will be discussed in detail in Section 2. In Section 3, we propose the underlying structures of the three types of complements and account for the semantic and syntactic behavior of these complements. More specifically, we propose that the *ko* complement in (4a) is a full CP, having the highest C head lexicalized by *ko* and the *kes* complements in (4b, c) are nominal in nature and structurally represent an NP, whose head is filled by *kes*. Yet, the two *kes* complements, with and without the overt tense and force markers, are further distinguished from each other: *kes* selects a full CP in (4b) and a reduced CP (without the projection of the highest C head) in (4c). Section 4 concludes the paper.

¹ The list of abbreviations used in this paper is the following: ACC-accusative, COMP-complementizer, DECL-declarative, INT-interrogative, LNK-linker, NEG-negation, NOM-nominative, TOP-topic.

2. The semantics and syntax of clausal complements in Korean

2.1. Presupposition

It is generally agreed in the literature that while the complement of a factive predicate is presupposed, that of a non-factive predicate is not (recall (1)). However, the following examples in Korean challenge such a dichotomous view.

- (5) a. Kibo-nun [Dana-ka i chayk-ul ilk-ess-ta-ko]
 Kibo-TOP Dana-NOM this book-ACC read-PAST-DECL-COMP
- b. Kibo-nun [Dana-ka i chayk-ul ilk-ess-ta-nun kes-ul]
 Kibo-TOP Dana-NOM this book-ACC read-PAST-DECL-LNK thing-ACC
- c. Kibo-nun [Dana-ka i chayk-ul ilk-un kes-ul]
 Kibo-TOP Dana-NOM this book-ACC read-LNK thing-ACC

yukamsulewehay-ss-ta / mit-ess-ta
 regret-PAST-DECL believe-PAST-DECL
 ‘Kibo regretted/believed that Dana read this book.’

kulente sasil-un Dana-nun i chayk-ul ilk-ci anh-ass-ta
 but fact-TOP Dana-TOP this book-ACC read-NEG-PAST-DECL

- a'. #regret ^{OK}believe
 b'. #regret ^{OK}believe
 c'. #regret #believe

(5) suggests that factivity and presupposition may not be as closely tied to each other as originally noted in Kiparsky & Kiparsky (1970). While it is true that the clausal complement of the factive predicate *yukamsuleweha* ‘regret’ is always presupposed in (5), it is not the case that the complement of the non-factive predicate *mit* ‘believe’ is never presupposed. This is formalized in (5a'-c'). In (5c'), we see that the *kes* complement is always presupposed whether it is selected by a factive predicate or a non-factive predicate.

Under the microscopic analysis, both the *ko* complement in (5a) and the *kes* complement with the overt past tense morpheme *-ss* and the declarative force marker *ta* in (5b) are unspecified for factivity. On the other hand, the *kes* complement without an overt tense morpheme and a force marker in (5c) is ALWAYS PRESUPPOSED regardless of the matrix predicate. Thus, it seems that the presuppositional reading of clausal complements is not solely determined by the nature of the matrix predicate, factive vs. non-factive, but derived from two related factors, namely the type of the matrix predicate (factives or non-factives) and the type of the complement in Korean (the *ko* clause, the *kes* clause with overt tense and force, and the *kes* clause without overt tense and force).

2.2. Wh-movement

To repeat from Section 1, research on factive and non-factive predicates in English shows that the clausal complement of the factive predicate is a weak island, out of which only object extraction is marginally allowed, whereas subject and adjunct extraction are not allowed. On the other hand, the clausal complement of the non-factive predicate does not exhibit any of these island constraints. To see whether similar observations are made in

Korean, three different types of *wh*-movement (objects, subjects, and adjuncts) were investigated.

However, it should be noted that Korean is a *wh-in-situ* language and therefore the existence of islandhood cannot be determined by overt dependency between two structural positions as in English. Yet, it has been convincingly argued in the literature that there is covert *wh*-movement at Logical Form (LF) in *wh-in-situ* languages, similarly to overt *wh*-movement in *wh-ex-situ* languages (Huang 1982b, Pesetsky 1987, Nishigauchi 1990, Watanabe 1992, Beck & Kim 1997, Ko 2005). Following this view, we assume that there is *wh*-movement in Korean at LF and argue that the (un)availability of *wh*-reading provides the evidence showing whether the complement forms an island to covert *wh*-movement.

2.2.1. Wh-objects

The examples in (6) show that *wh*-object questions are formed out of the three sentential complements under factive and non-factive predicates, (a) the *ko* clause, (b) the *kes* clause with overt tense and force, and (c) the *kes* clause without overt tense and force.

- (6) a. Kibo-nun [Dana-ka mwuet-ul ilk-ess-ta-ko]
 Kibo-TOP Dana-NOM what-ACC read-PAST-DECL-COMP
- b. Kibo-nun [Dana-ka mwuet-ul chayk-ul ilk-ess-ta-nun kes-ul]
 Kibo-TOP Dana-NOM what-ACC book-ACC read-PAST-DECL-LNK thing-ACC
- c. Kibo-nun [Dana-ka mwuet-ul chayk-ul ilk-un kes-ul]
 Kibo-TOP Dana-NOM what-ACC book-ACC read-LNK thing-ACC

yukamsulewehay-ss-ni ? / mit-ess-ni
 regret-PAST-INT believe-PAST-INT
 ‘What did Kibo regret/believe that Dana read ___?’

- a. ^{OK}regret ^{OK}believe
 b. *regret *believe
 c. *regret *believe

We observe that regardless of the matrix predicate, whether it is factive or non-factive, covert *wh*-object movement is only possible out of the *ko* complement (6a), but not out of the *kes* complement (6b, c). This tells us that the *ko* complement is not an island to *wh*-object movement, but the *kes* complements form an island to covert *wh*-object movement in Korean. While it seems that the type of the complement (*ko* vs. *kes*) is a determining factor allowing *wh*-object movement, the choice of the matrix predicate also plays a role in the case of *wh*-subject and *wh*-adjunct movement, as will be discussed below.

2.2.2. Wh-subjects and wh-adjuncts

(7) and (8) illustrate covert *wh*-subject and *wh*-adjunct movement out of the three sentential complements under factive and non-factive predicates in Korean. Both types of *wh*-movement exhibit a similar, yet more restricted pattern than that of *wh*-object movement.

- (7) a. Kibo-nun [nwu-ka i chayk-ul ilk-ess-ta-ko]
 Kibo-TOP who-NOM this book-ACC read-PAST-DECL-COMP
- b. Kibo-nun [nwu-ka i chayk-ul ilk-ess-ta-nun kes-ul]
 Kibo-TOP who-NOM this book-ACC read-PAST-DECL-LNK thing-ACC
- c. Kibo-nun [nwu-ka i chayk-ul ilk-un kes-ul]
 Kibo-TOP who-NOM this book-ACC read-LNK thing-ACC
- yukamsulewehay-ss-ni ? / mit-ess-ni
 regret-PAST-INT believe-PAST-INT
 ‘Who did Kibo regret/believe (that) ___ read this book?’
- a. **regret* ^{OK}*believe*
 b. **regret* **believe*
 c. **regret* **believe*
- (8) a. Kibo-nun [Dana-ka way i chayk-ul ilk-ess-ta-ko]
 Kibo-TOP Dana-NOM why this book-ACC read-PAST-DECL-COMP
- b. Kibo-nun [Dana-ka way i chayk-ul ilk-ess-ta-nun kes-ul]
 Kibo-TOP Dana-NOM why this book-ACC read-PAST-DECL-LNK thing-ACC
- c. Kibo-nun [Dana-ka way i chayk-ul ilk-un kes-ul]
 Kibo-TOP Dana-NOM why this book-ACC read-LNK thing-ACC
- yukamsulewehay-ss-ni ? / mit-ess-ni
 regret-PAST-INT believe-PAST-INT
 ‘Why did Kibo regret/believe that Dana read this book ___?’
- a. **regret* ^{OK}*believe*
 b. **regret* **believe*
 c. **regret* **believe*

Similar to *wh*-object movement in (6), the two types of *kes* complements, with and without overt tense and force markers (7b, c) and (8b c), are an island to both *wh*-subject and *wh*-adjunct movement in Korean. However, the *ko* complement exhibits a slightly different pattern: while the *ko* complement under the factive predicate *yukamsuleweha* ‘regret’ is an island for *wh*-subject and *wh*-adjunct movement, the *ko* complement is not an island when it is selected by the non-factive predicate *mit* ‘believe’ (7a, 8a). Thus, it seems that both the complement type (*ko* vs. *kes*) and the matrix predicate (factive vs. non-factive) play a role to determine the islandhood of the clausal complement for non-complement *wh*-movement.

2.3. Scrambling

In Section 2.2, we examined *wh*-movement out of the three sentential complements in Korean. Since there is no overt *wh*-movement in this language, to argue for the islandhood of these complements for *wh*-movement, we assumed that there is covert *wh*-phrasal movement in Korean, following the proposals made in the literature. Yet, to bolster the claim that a

subset of these clausal complements is opaque to movement, we further investigate the phenomenon of SCRAMBLING, which occurs in overt syntax.

There are two different views on the phenomenon of scrambling in the literature, a movement approach (Ross 1967, Saito 1985, 1992, Mahajan 1994, Müller 1996, Kang 2005, among others) and a base generation approach (Bayer & Kornfilt 1994, Miyagawa 1997, 2001, Fanselow 2001). However, scrambling in Korean is convincingly argued to involve movement rather than being base-generated (Kang 1994, Lee 2007), the view we adopt in this paper. In this section we concentrate on object scrambling to test the presence vs. absence of the islandhood for the three types of clausal complements in Korean. The reason why we limit ourselves to object scrambling only is that the nature of subject and adjunct scrambling is rather poorly understood compared to object scrambling. For instance, Saito (1985) argues that subject scrambling, both clause internally and externally, is impossible in Japanese based on the contrast found in object vs. subject scrambling. While the same empirical facts that Saito discusses to argue against subject scrambling also hold in Korean, there are counter-examples provided in subsequent work (Miyagawa 1989 for Japanese, Sohn 1995 and Ko 2008 for Korean), which seems to support the existence of subject scrambling. Skirting away from the debate on subject and adjunct scrambling in the literature, we therefore focus on object scrambling in the following.

- (9) a. i chayk_i-ul Kibo-nun [Dana-ka t_i ilk-ess-ta-ko]
 this book-ACC Kibo-TOP Dana-NOM read-PAST-DECL-COMP
- b. i chayk_i-ul Kibo-nun [Dana-ka t_i ilk-ess-ta-nun kes-ul]
 this book-ACC Kibo-TOP Dana-NOM read-PAST-DECL-LNK thing-ACC
- c. i chayk_i-ul Kibo-nun [Dana-ka t_i ilk-un kes-ul]
 this book-ACC Kibo-TOP Dana-NOM read-LNK thing-ACC
- yukamsulewehay-ss-ta / mit-ess-ta
 regret-PAST-DECL believe-PAST-DECL
 ‘This book, Kibo regretted/believed that Dana read ____.’
- a. ^{OK}regret ^{OK}believe
 b. ^{OK}regret ^{OK}believe
 c. *regret *believe

In (9), the object can be scrambled out of the *ko* complement (9a) and the *kes* complement with the overt tense morpheme and the force marker (9b). On the other hand, when the *kes* complement is bare, without tense and force markers (9c), object scrambling is prohibited or much more degraded at best. Interestingly, object scrambling in overt syntax does not pattern together with *wh*-object movement at LF in (6), in which only the *ko* complement is transparent to *wh*-object movement. The Table 1 below summarizes all the aforementioned semantic and syntactic behavior of the three types of clausal complements in Korean.

Table 1: The interplay between the matrix predicate and its clausal complement

		Types of matrix predicates	
		<i>regret</i> (factive)	<i>believe</i> (non-factive)
Types of complements	(a) <i>ko</i>	presupposition OK object scrambling OK <i>wh</i> -object * <i>wh</i> -subject * <i>wh</i> -adjunct	no presupposition OK object scrambling OK <i>wh</i> -object OK <i>wh</i> -subject OK <i>wh</i> -adjunct
	(b) <i>kes</i> with overt tense and force	presupposition OK object scrambling * <i>wh</i> -object * <i>wh</i> -subject * <i>wh</i> -adjunct	no presupposition OK object scrambling * <i>wh</i> -object * <i>wh</i> -subject * <i>wh</i> -adjunct
	(c) <i>kes</i> without overt tense and force	presupposition * object scrambling * <i>wh</i> -object * <i>wh</i> -subject * <i>wh</i> -adjunct	presupposition * object scrambling * <i>wh</i> -object * <i>wh</i> -subject * <i>wh</i> -adjunct

3. Proposal and analysis

To account for the empirical facts (presupposition and islandhood) provided in Table 1, we propose the following underlying structures for the three sentential complements in Korean.²

- (10) a. *ko* complements [_{CP} *ko* [_{ForceP} *ta* [_{TP}]]]
 b. *kes* complements with overt tense and force [_{NP} *kes* [_{CP} \emptyset_{KO} [_{ForceP} *ta* [_{TP}]]]]
 c. *kes* complements without overt tense and force [_{NP} *kes* [_{ForceP} \emptyset [_{TP}]]]

3.1. *Ko* complements

Table 1 shows that the *ko* complements selected by a factive predicate and a non-factive predicate behave differently. While the *ko* complement under the factive predicate is presupposed, the one under the non-factive predicate is not. Also, the former is a weak island to *wh*-movement, allowing only object movement, the latter is not an island to *wh*-movement at all; *wh*-object, *wh*-subject and *wh*-adjunct movement are all possible out of the *ko* complement under the non-factive predicate. The contrast found in the *ko* complement embedded under factive and non-factive predicates in Korean in fact mimics the contrast between factive and non-factive *that* complements in English; factive *that* complements are presupposed and form a weak island to *wh*-movement, whereas non-factive complements are not presupposed and not an island.

We propose that the complementizer *ko* lexicalizes the highest C head in a split CP structure (Rizzi 1997), and it is unspecified for factivity. Thus, the presupposition of the *ko* complement is determined by the nature of the matrix predicate. In other words, the *ko*

² In this paper, we limit ourselves to provide an account of the (un)availability of presuppositional reading of the clausal complement and the patterns of object scrambling and *wh*-object movement out of the clausal complement. Further discussion of subject scrambling/extraction and *wh*-subject and *wh*-adjunct movement is provided in Shim & Ihsane (forthcoming).

complement of the factive predicate is presupposed, and the *ko* complement of the non-factive predicate is not presupposed, similar to English. (11) represents the underlying structure of the *ko* complement, repeated from (10a).³

(11) [CP *ko* [_{ForceP} *ta* [TP]]]

The structure in (11) also accounts for both overt and covert object movement out of the *ko* complement. Assuming that both overt scrambling and covert *wh*-movement are subject to locality constraints, the object first moves up to the Spec CP in the embedded clause and then to the matrix clause either in overt syntax or at LF, resulting in object scrambling and *wh*-object movement out of the *ko* complement, respectively.

In contrast to *wh*-object movement, which is allowed out of the *ko* complement under both factive and non-factive predicates alike, non-complement *wh*-movement, such as subject and adjunct movement, exhibits a more restricted pattern. While both *wh*-subject and *wh*-adjunct movement are possible out of the *ko* complement under the non-factive predicate *believe*, they are banned out of the *ko* complement under the factive predicate *regret*. However, this contrast is only found in the *ko* complement, and both *wh*-subject and *wh*-adjunct movement are forbidden out of the *kes* complements regardless of the matrix predicate, as shown in Table 1. In other words, non-complement *wh*-movement out of the embedded clause (subjects and adjuncts) is more restricted than complement *wh*-movement in Korean.

To account for the more limited patterns of *wh*-subject and *wh*-adjunct movement out of the clausal complement in Korean, we follow the prevalent views on subject/adjunct vs. object asymmetry in the literature, such as the Empty Category Principle (ECP; Chomsky 1981) and the Condition on Extraction Domains (CED; Huang 1982a).⁴ In the Government and Binding (GB) theory, the ECP states that a trace must be properly governed, either lexically or antecedent governed. The fact that covert *wh*-subject and *wh*-adjunct movement are prohibited out of the clausal complement in Korean can be explained by the idea that subject and adjunct movement at LF result in the configuration in which their traces are not properly governed. While subject movement and adjunct movement in overt syntax allow the moved subject and adjunct to c-command and thus antecedent govern their traces, LF movement of the subject and the adjunct does not lead to a position where the subject and the adjunct can c-command their traces, thus the traces cannot be properly (antecedent) governed. On the other hand, *wh*-object movement at LF is licit in Korean, for the trace can be lexically governed by the verb, therefore properly governed.

³ The structure in (11) involves multiple CP layers in a split CP structure, in which there is a projection above ForceP. In Rizzi (1997)'s original work, ForceP is the highest projection in the multi-layered structure for complementizers. However, ensuing cross-linguistic studies on complementizers show that there is an additional layer above ForceP in the complementizer structure, which may be morphologically marked above the complementizer (e.g., *hai* in Romanian (Hill 2006); *diz* in American Spanish (Demonte & Fernández-Soriano 2013)). In addition, it is argued that the additional layer above ForceP is related to syntax-pragmatics interface, such as evidentiality and reported speech. Similar analyses have been offered for the complementizer *ko* in Korean, relating it to evidentiality (Ahn & Yap 2012) or reported speech (Chang 1996), which justifies the structure proposed in (11).

In Shim & Ihsane (2015), we provide a different analysis of the complementizer *ko*, in which *ko* is re-analyzed as a light D head taking a CP complement, a structure similar to Hungarian *azt-hogy* constructions.

⁴ At first glance, it does not seem that *wh*-subject and *wh*-adjunct movement are more restricted than *wh*-object movement out of the clausal complement in Korean since *wh*-object movement is categorically banned out of the two types of *kes* complements as well, as shown in Table 1. However, we will show in Sections 3.2. and 3.3. that illicit *wh*-object movement out of the *kes* complements is due to the locality constraints, not due to its limited nature unlike non-complement *wh*-movement, and maintain the view on subject/adjunct vs. object asymmetries.

While the ECP-based explanation provided above accounts for the general restrictions on covert *wh*-subject and *wh*-adjunct movement out of the clausal complements in Korean, it does not explain why *wh*-subject and *wh*-adjunct movement are possible out of the *ko* complement under the non-factive predicate in Table 1. For the moment, we have no account for this and leave this for future research.⁵

3.2. *Kes* complements

The clausal complements headed by *kes* are divided into two sub-types, one with an overt tense morpheme (e.g., *-ess* PAST) and the declarative force marker *ta* and the other without an overt tense morpheme or a force marker, as shown in (4b) and (4c) respectively, which are repeated in (12).

- (12) a. Kibo-nun [Dana-ka i chayk-ul ilk-**ess-ta**-nun **kes-ul**]
 Kibo-TOP Dana-NOM this book-ACC read-PAST-DECL-LNK thing-ACC
- b. Kibo-nun [Dana-ka i chayk-ul ilk-un **kes-ul**]
 Kibo-TOP Dana-NOM this book-ACC read-LNK thing-ACC
- yukamsulewehay-ss-ta / mit-ess-ta
 regret-PAST-DECL believe-PAST-DECL
 ‘Kibo regretted/believed that Dana read this book.’

The word *kes* is roughly translated into ‘(a) thing’ in English and it has been analyzed as a bound noun (Sohn & Nam 2013) or a nominalizer (Jhang 1994, Chung 1999, Kim, M.-J. 2004). Following the converging views of treating *kes* as a nominal element, we claim that *kes* lexicalizes an N head, thus the phrase headed by *kes* is nominal in nature. Based on this, we propose the following underlying structures of the two types of *kes* complements in (13).

- (13) a. *kes* complements with overt tense and force [NP *kes* [CP \emptyset_{KO} [ForceP *ta* [TP T]]]]
 b. *kes* complements without overt tense and force [NP *kes* [ForceP \emptyset [TP]]]

In (13a), *kes* selects a CP headed by a null allomorph of the complementizer *ko*. The evidence for postulating the null complementizer *ko* in (13a) comes from the fact that *ko* can be in fact optional in the *ko* complement in (11). Adopting the proposal that null complementizers are affixes (Pesetsky 1992), we assume that the null complementizer *ko* is an affix and propose that it incorporates into the N head *kes*, forming a complex head at LF.⁶

⁵ Another potential problem of the ECP account of subject/adjunct vs. object asymmetries is that the notion of GOVERNMENT, which played an essential role in the GB theory, can no longer be formulated in the Minimalist Program, the framework we adopt in this paper. There have been proposals made to capture the systematic patterns of subject/adjunct vs. object asymmetries in the more recent Minimalist literature (e.g., Takahashi 1994, Nunes & Uriagereka 2000). As pointed out by Stepanov (2007), however, these proposals fail to account for a range of empirical facts. We intend to investigate this further in the future.

⁶ If our proposal is right that there is a null C heading the complement of the noun in the *kes* complement in (13a), it seems that languages differ with respect to the environment licensing a null complementizer. While the null C can be licensed by its selecting N head in Korean, it is a well-known fact that the null *that* is not allowed in sentential complementation to a noun in English (e.g., *Kibo regretted/believed the news *(that) Dana read this book*).

Such cross-linguistic variation on licensing a null complementizer can be accounted for by the views that affixes have subcategorization requirements, and the null C may selectively take a lexical head as its host

The presupposition of the *kes* complement with overt tense and force in (13a) is then determined by the interaction between the matrix predicate, factive or non-factive, and the complex C+N head (\emptyset_{KO+kes}) unspecified for factivity. As a result, it is determined by the nature of the matrix predicate, similar to that of the *ko* complement.

On the other hand, in the *kes* complement without overt tense and force in (13b), *kes* does not select a fully projected CP but a reduced CP structure. There is no projection of the highest C headed by the null complementizer *ko*, but a ForceP. It is important to note that the Force head in (13b) cannot be occupied by the null declarative marker *ta*, but remains empty. Unlike the complementizer *ko*, the declarative marker *ta* does not have a null allomorph, for it can never occur optionally, as shown in (14).

- (14) a. Kibo-nun [Dana-ka i chayk-ul ilk-ess-*(**ta**)-ko]
 Kibo-TOP Dana-NOM this book-ACC read-PAST-DECL-COMP
- b. Kibo-nun [Dana-ka i chayk-ul ilk-ess-*(**ta**)-nun kes-ul]
 Kibo-TOP Dana-NOM this book-ACC read-PAST-DECL-LNK thing-ACC
- yukamsulewehay-ss-*(**ta**) / mit-ess-*(**ta**)
 regret-PAST-DECL believe-PAST-DECL
 ‘Kibo regretted/believed that Dana read this book.’

One may ask why the structure in (13b) includes both a ForceP and a TP even though there is no morphological evidence marking Force and Tense in the *kes* complement of the type in (12b); the verb *ilk* is in the bare form. The projection of Force and T in (13b) comes from the Minimalist views that (i) nominative Case is a result of agreement/feature matching between the probe T and the goal (the external argument/subject), and (ii) all of T’s features, such as tense, ϕ -, and (nominative-) Case features, are inherited from the C head above T (FEATURE INHERITANCE à la Chomsky 2008). Thus, the fact that the embedded subject *Dana* in (12b) is nominative Case-marked suggests that there are T and C nodes projected in the embedded clause, and the nominative Case on the embedded subject *Dana* is licensed via agreement with Force-T.

One important element that should be noted here is that in Chomsky’s system, there is only one C head above TP, and T inherits its features from the C. However, we propose a multi-layered CP structure and assume that T may inherit its features from multiple C heads. What feature inheritance suggests is that there is an additional functional C-type category above a (finite) TP, from which T inherits its features, and we propose that it is Force(P) in (13b).

Yet, the structures in (13a) and (13b) differ from each other. While the Force head is lexicalized by *ta* in the former, it is radically empty in the latter. We interpret this in such a way that the empty Force head in (13b) does not carry an ‘assertion’ feature, in contrast with the Force head in (13a). Basse (2008) argues that the presupposition of factive complements in English stems from the lack of an assertion feature (relative to the matrix subject) in ForceP in the clausal complement, and the absence of matrix subject assertion causes the embedded proposition to anchor to the speaker and is therefore interpreted as a speaker presupposition. Similarly, we propose that, due to the absence of the assertion feature in ForceP in (13b), presuppositional reading arises as a default, and that the embedded proposition is not being asserted or related to the speaker but is taken for granted (cf. Hooper & Thompson 1973,

(Bošković & Lasnik 2003); while the null C can be hosted by [+V] elements in English, it can be hosted by either [+V] or [+N] elements in Korean.

Haegeman 2006). This accounts for the fact that the *kes* complement without the overt tense and force markers in (12b) is always presupposed regardless of the matrix predicate, whether it is factive or non-factive, as summarized in Table 1 above.

To make our claim more concrete, let us discuss in detail the example given in (5c), in which the bare *kes* complement is embedded under a factive and a non-factive complement, which is repeated below.

- (15) a. Kibo-nur [Dana-ka i chayk-ul ilk-un kes-ul]
Kibo-TOF Dana-NOM this book-ACC read-LNK thing-ACC

yukamsulewehay-ss-ta / mit-ess-ta
regret-PAST-DECL believe-PAST-DECL

kulente sasil-un Dana-nun i chayk-ul ilk-ci anh-ass-ta
but fact-TOP Dana-TOP this book-ACC read-NEG-PAST-DECL

‘Kibo regretted/believed that Dana read this book, but in fact she didn’t read it.’

#*regret* #*believe*

- b. Kibo-nur [Dana-ka i chayk-ul ilk-un kes-ul]
Kibo-TOF Dana-NOM this book-ACC read-LNK thing-ACC

yukamsuleweha-cianh-ass-ta / mit-ciahn-ass-ta
regret-NEG-PAST-DECL believe-NEG-PAST-DECL

kulente sasil-un Dana-nun i chayk-ul ilk-ci anh-ass-ta
but fact-TOP Dana-TOP this book-ACC read-NEG-PAST-DECL

‘Kibo did not regret/believe that Dana read this book, but in fact she didn’t read it.’

#*regret* #*believe*

In (15a), the presuppositional reading of the *kes* complement, whether it is selected by a factive predicate or a non-factive predicate, stems from the lack of an assertion feature (relative to the matrix subject) in ForceP, whose head is radically empty. In other words, both the speaker and the matrix subject *Kibo* presuppose the truth of the complement clause, *Dana read this book*. We further notice that the presupposition of the *kes* complement without overt tense and force remains constant under the scope of matrix negation in (15b), which is a known property of determining presuppositionality (à la Kiparsky & Kiparsky 1970).

In addition to presuppositionality, the two types of *kes* complements, with and without overt tense and force, exhibit further differences. While object scrambling is possible in (13a), it is forbidden in (13b). But covert *wh*-movements (object, subject, and adjuncts) out of both types of *kes* complements are banned, as Table 1 illustrates. Why is then object movement possible in overt syntax but not at LF in (13a)? The embedded object can undergo scrambling via Spec, CP, as the head C is realized by the null complementizer *ko* (\emptyset_{KO}). However, this Spec, CP position is no longer available for covert *wh*-object movement due to PHASE EXTENSION (16):

- (16) Phase Extension

Syntactic movement of the head H of a phase α up to the X of a node β dominating α extends the phase up from α to β ; α loses its phasehood in the process, and any

constituent on the edge of α ends up in the domain of the derived phase β as result of phase extension (den Dikken 2006: 1)

After the null C head is incorporated into N at LF, the CP phasehood is extended to the NP headed by *kes*. Given the PIC (Phase Impenetrability Condition; Chomsky 2000), the *wh*-object would need to move through Spec, NP after phase extension from the CP to the NP. This movement, however, is prohibited since Spec, NP is either an A-position or is not licensed unless N theta-marks the element in this position (Bošković 2008). Thus, C-to-N incorporation at LF blocks all covert *wh*-movement out of the *kes* complement in (13a).

On the other hand, in (13b), there is no Spec, CP available for the object to move through either in overt syntax or at LF. Here we assume that only the highest C head is a phase head in the split CP structure as in C-Force and only the phase head and its edge are accessible to syntactic operations. Thus, in the reduced CP structure in (13b), in which the phase head C is not projected, all syntactic operations/movement are not allowed.

4. Conclusion

This paper is against the simplistic dichotomous views on factive vs. non-factive complements, in which factive complements are always presupposed and are a weak island to *wh*-movement, whereas non-factive complements are never presupposed and do not form an island. Detailed analyses of three different types of clausal complements in Korean reveal that the diverse semantic and syntactic behavior (factivity and islandhood) of clausal complements is determined by the interplay between two related factors, the type of the matrix predicate (factive vs. non-factive) and the type of the clausal complement.

We have argued that the three types of clausal complements in Korean have different underlying structures; (i) the *ko* complement is a CP, (ii) the *kes* complement with overt tense and force is an NP taking a full CP, and (iii) the *kes* complement without overt tense and force is an NP taking a reduced CP. Based on these, we have provided a minimalist account of the intricate semantic and syntactic patterns emerging out of these complements.

If the present analysis is on the right track, it has a few important implications. First, it shows that the complementizer itself is not the locus of presupposition and factivity in Korean (*pace* Yoon 2013). Instead, presupposition arises from the interplay between the matrix predicate and the type of the complement. Our proposal also provides a way to analyze *wh-in-situ* languages on a par with *wh-ex-situ* languages by arguing that they are both subject to the same locality constraints. Thus, one need not postulate different assumptions or rules distinguishing between overt and covert syntactic movement/operations.

References

- Ahn, M. & Yap, F. H. (2012). On the extended uses of evidential markers in Korean: diachronic and discourse perspectives. Paper presented at the Conference on the Nature of Evidentiality. Leiden: Netherlands.
- Basse, G. (2008). Factive complements as defective phases. *Proceedings of the 27th West Coast Conference on Formal Linguistics*, 54-62.
- Bayer, J. & Kornfilt, J. (1994). Against scrambling as an instance of move-alpha. In Corver, N., & van Riemsdijk, H. (eds.) *Studies on scrambling: movement and non-movement approaches to free word-order phenomena*, 17-60. Berlin: Mouton de Gruyter.
- Beck, S. & Kim, S.-S. (1997). On *wh*- and operator scope in Korean. *Journal of East Asian Linguistics* 6, 339-384.

- Bošković, Z. (2008). What will you have, DP or NP? *Proceedings of North East Linguistic Society* 37, 101-114.
- Bošković, Z. & Lasnik, H. (2003). On the distribution of null complementizers. *Linguistic Inquiry* 34, 527-546.
- Cattell, R. (1978). On the source of interrogative adverbs. *Language* 54, 61-77.
- Chang, S.-J. (1996). *Korean*. Amsterdam, Netherlands: John Benjamins.
- Chomsky, N. (1981). *Lectures on Government and Binding: the Pisa lectures*. Holland: Foris Publications.
- Chomsky, N. (2000). Minimalist inquiries: the framework. In Martin, R., Michaels, D., & Uriagereka, J. (eds.) *Step by step: essays on minimalist syntax in honor of Howard Lasnik*, 89-156. Cambridge, MA: MIT Press.
- Chomsky, N. (2001). Derivation by phase. In Kenstowicz, M. (ed.) *Ken Hale: a life in language*, 1-52. Cambridge, MA: MIT Press.
- Chung, D.-H. (1999). A complement analysis of the head internal relative clauses. *Language and Information* 3, 1-12.
- De Cuba, C. & Ürögdi, B. (2001). The role of factivity in the syntax and semantics of embedded clauses. Ms., Stonybrook University.
- Demonte, V. & Fernández-Soriano, O. (2013). Evidentials *dizque* and *que* in Spanish: Grammaticalization, parameters and the (fine) structure of Comp. *Linguística. Revista de Estudos linguísticos da Universidade do Porto* 8, 211-234.
- Den Dikken, M. (2006). Phase extension: contours of a theory of the role of head movement in phrasal extraction. Ms., City University of New York.
- Fanselow, G. (2001). Features, θ -roles, and free constituent order. *Linguistic Inquiry* 32, 405-436.
- Haegeman, L. (2006). Conditionals, factives and left periphery. *Lingua* 116, 1651-1669.
- Hegarty, M. (1990). On adjunct extraction from complements. *MIT Working Papers in Linguistics* 13, 101-124.
- Hegarty, M. (1992). Familiar complements and their complementizers: on some determinants of A'-locality. Ms., University of Pennsylvania.
- Hill, V. (2006). Pragmatic markers as syntactic heads: a case study from Romanian. In Grohmann, K. K., & Panagiotidis, P. (eds.) *Selected papers from the 2006 Cyprus syntaxfest*, 237-265.
- Hooper, J. B. & Thomson, S. A. (1973). Transformations. *Linguistic Inquiry* 4, 465-497.
- Huang, C.-T. J. (1982a). Logical relations in Chinese and the theory of grammar. Ph.D. dissertation, Massachusetts Institute of Technology.
- Huang, C.-T. J. (1982b). Move WH in a language without WH movement. *The Linguistic Review* 1, 369-416.
- Jhang, S. (1994). *Headed nominalizations in Korean: relative clauses, clefts, and comparatives*. Ph.D. dissertation, Simon Fraser University.
- Kang, D.-H. (2005). *Scrambling in Universal Grammar: an analysis of scrambling as optional movement in Korean and other languages*. Ph.D dissertation, University of Stuttgart.
- Kang, Y. E. (1994). *Weak and strong interpretations of quantifiers and definite NPs in English and Korean*. Ph.D. dissertation, University of Texas at Austin.
- Kim, M.-J. (2004). Three types of *kes*-nominalization in Korean. *Harvard Studies in Korean Linguistics* 10, 479-492.
- Kiparsky, P. & Kiparsky, C. (1970). Fact. In Steinberg, D., & Jakobovits, L. (eds.) *Semantics*, 345-369. Cambridge, UK: Cambridge University Press.
- Ko, H. (2005). Syntax of *wh*-in-situ: merge into [Spec, CP] in the overt syntax. *Natural Language & Linguistic Theory* 23, 867-916.

- Ko, H. (2008). Subject scrambling. *Japanese/Korean Linguistics* 13.
- Lee, E. (2007). *Types of scrambling in Korean syntax*. Ph.D dissertation, University of North Carolina at Chapel Hill.
- Mahajan, A. (1994). The ergativity parameter: Have-be alternation, word order and split ergativity. *Proceedings of North East Linguistic Society* 24, 317-331.
- Melvold, J. (1991). Factivity and definiteness. *MIT Working Papers in Linguistics* 5, 97-117.
- Miyagawa, S. (1989). *Structure and case marking in Japanese*. San Diego, CA: Academic Press.
- Miyagawa, S. (1997). Against optional scrambling. *Linguistic Inquiry* 28, 1-25.
- Miyagawa, S. (2001). Agreements that occur mainly in the main clause. Ms., Massachusetts Institute of Technology.
- Müller, G. (1996). A constraint on remnant movement. *Natural Language & Linguistic Theory* 14, 355-407.
- Nishigauchi, T. (1990). *Quantification in the theory of grammar*. Dordrecht, Boston, London: Kluwer Academic Publishers.
- Nunes, J. & Uriagereka, J. (2000). Cyclicity and extraction domains. *Syntax* 3, 20-43.
- Pesetsky, D. (1987). Wh-in-situ: movement and unselective binding. In Reuland, E. & ter Meulen, A. (eds.) *The representation of (in)definiteness*, 98-129. Cambridge, MA: MIT Press.
- Pesetsky, D. (1992). Zero syntax Vol. 2. Ms., Massachusetts Institute of Technology.
- Rizzi, L. (1997). The fine structure of left periphery. Ms., University of Sienna.
- Ross, J. R. (1967). *Constraints in variables in syntax*. Ph.D dissertation, Massachusetts Institute of Technology.
- Saito, M. (1985). *Some asymmetries in Japanese and their theoretical implications*. Ph.D dissertation, Massachusetts Institute of Technology.
- Saito, M. (1992). Long distance scrambling in Japanese. *Journal of East Asian Linguistics* 1, 69-118.
- Shim, J. Y. & Ihsane, T. (2015). The unbearable lightness of being a ϕ -incomplete D. Ms., University of Geneva.
- Sohn, K.-W. (1995). *Negative polarity items, scope and economy*. Ph.D dissertation, University of Connecticut at Storrs.
- Sohn, S. O. & Nam, K. (2013). The formulaic expressions of ‘thing’ in spoken and written registers: a corpus study of bound nouns *kes* and *ke* in Korean. *Japanese/Korean Linguistics* 20, 31-48.
- Stepanov, A. (2007). The end of CED? Minimalism and extraction domains. *Syntax* 10, 80-126.
- Takahashi, D. (1994). *Minimality of movement*. Ph.D dissertation, University of Connecticut at Storrs.
- Varlakosta, S. (1994). Factivity, factive complements and wh-extraction. *MIT Working Papers in Linguistics* 23, 317-336.
- Watanabe, A. (1992). Subjacency and S-structure movement of wh-in-situ. *Journal of East Asian Linguistics* 1, 255-291.
- Yoon, S. (2013). On embedded room phenomena: ERP in Korean versus V2 in Germanic languages. *Proceedings of North East Linguistic Society* 42.

*Ji Young Shim
Département de Linguistique
Faculté des Lettres
Université de Genève
CH-1211 Genève 4
Suisse*

Jiyoung.shim@unige.ch

*Tabea Ihsane
Département de Linguistique
Faculté des Lettres
Université de Genève
CH-1211 Genève 4
Suisse*

Tabea.ihsane@unige.ch

WH-SCOPE MARKING IN SYRIAN ARABIC: AN INDIRECT DEPENDENCY

MAIS SULAIMAN
(Newcastle University)

Abstract

This paper discusses wh-scope marking in Syrian Arabic and approaches to wh-scope marking. It is argued that Syrian Arabic employs the wh-scope marking strategy for questioning out of embedded questions. The dependency between the wh-scope marker and the embedded clause is indirect. The wh-scope marker is base generated in A-position and moves to Spec of the matrix clause. The embedded clause is an unselected complement of the verb. This analysis supports Dayal's (2000) proposal that a unified theory of wh-scope marking cross-linguistically can be reconciled under the indirect-dependency analysis.

1. Introduction

Syrian Arabic (SA) employs long wh-movement for questioning out of embedded questions. Wh-scope marking or partial wh-movement can also be employed as an alternative strategy, as illustrated in (1):

- (1) a. maʕ miin fkaar-ti-ni knt ʕam iħk-i?
with who thought-2SG.F.SU-1SG.OBJ was.1SG PROG speaking
'Who did you think I was talking to?'
- b. ŕu fkkar-ty maʕ miin knt ʕam iħk-i?
what thought-2SG.F.SU with who was.1SG PROG speaking
'What did you think? Who was I talking to?'

In the long extraction question (1a), the wh-phrase *miin* 'who' undergoes long movement to Spec of the matrix clause. In the partial wh-movement question (1b), the wh-phrase raises to Spec of the embedded clause. The scope is marked by another wh-phrase *ŕu* 'what' in Spec of the matrix clause.

Based on the following evidence, it can be argued that sentences like (1b) are real instances of wh-scope marking, rather than a sequence of two independent sentences, each involving full wh-movement: wh-scope marking constructions in Syrian can occur in embedded contexts, as in (2):

- (2) a. sual-ek ŕw raʔy-w min raħ yntSer b-l-akhirma kan
question-your what opinion-his who will win with-the-end not was
fi daʕi il-w.
in reason for-it
'Your question about what he thinks who will win at the end was not appropriate.'
- b. ma tʔul-i-l-u ŕw ʔal-et-l-ek mama min
not say-2SG.F.SU-to-3SG.M.OBJ what said.3SG.F.SU-to-2SG.F.OBJ mom who
jayeh la-ʕanna bukra?
coming to-ours tomorrow
'Don't tell him what mom told you about who is visiting us tomorrow.'

Another piece of evidence comes from the fact that a pronoun in the embedded clause can be interpreted as a bound pronoun while its antecedent is in the first clause (see (3)). This entails that the two *wh*-clauses cannot be separate sentences. The pronoun *w* must be c-commanded by the antecedent *kl waḥed* ‘every one’.

- (3) šw fakkar kl-waḥed addesh raḥ yTlaḡ-l-w?
 what thought every-one how.much will get-to-3SG.M
 ‘How much does every one think he will get?’

Wh-scope marking is discussed in the literature in terms of two main theories: *the direct dependency* and *the indirect dependency*. It is argued that German is best interpreted in terms of a direct dependency analysis (Riemsdijk 1983; McDaniel 1989), while Hindi is best interpreted in terms of an indirect dependency (Dayal 1994, 2000; Beck & Berman 2000). Horvath (1997, 2000) argues that *wh*-scope marking varies cross-linguistically, thus each language requires a different interpretation.

In this paper, I argue that *wh*-scope marking in Syrian Arabic manifests an indirect dependency between the *wh*-scope marker and the embedded clause. The *wh*-scope marker is a theta-marked argument of the matrix verb base generated in the direct object position. I argue following Felser (2001) that *wh*-scope marking in SA manifests a predication relation between the *wh*-scope marker and the embedded *wh*-clause. However, this can be regarded as an identificational relation. The *wh*-clause provides an identification or definition of the *wh*-scope marker.

2. Properties of *wh*-scope marking in Syrian Arabic

Wh-scope marking in SA manifests properties of *wh*-scope marking constructions. I present some of these properties following Dayal (1994):

A. Any *wh*-phrase can be associated with the *wh*-scope marker, as in (4):

- (4) a. šw fkkart wen raʔset maria?
 what thought.2SG.M where danced Maria
 ‘What did you think? Where did Maria dance?’
- b. šw fkkart aymat raʔset maria?
 what thought.2GS.M when danced Maria
 ‘What did you think? When did Maria dance?’
- c. šw fkkart leš raʔset maria?
 what thought.2SG.M why danced Maria
 ‘What did you think? Why did Maria dance?’
- d. šw fkkart kif raʔset maria?
 what thought.2SG.M how danced Maria
 ‘What did you think? How did Maria dance?’

B. SA allows embedded yes/no questions in scope marking constructions.

- (5) a. šw al-l-ek Iyad bd-w yji?
 what said.3SG.M.SU-to-2SG.F.OBJ Iyad want-3SG.M come
 ‘What did Iyad say? Is he coming?’

- b. šw fkkar-ti ʕam iħk-i la-ħaly?
 what thought-2SG.F PROG speaking-1SG to-myself
 ‘Did you think I am talking to myself?’

Yes/ no questions in Syrian are marked by intonation only. They are not marked by any of the common question formation mechanisms, such as inversion or question particles that is wh-phrase equivalent.

C. Wh-scope marking can occur with multiple embedded wh-phrases associated with the scope marker. Multiple wh-questions in SA are allowed in discourse-linked contexts. A coordinative head appears before adverbial wh-phrases, as illustrated in (6):

- (6) a. šw fkkar-ty aymat rah basem w la-wen?
 what thought-2SG.F when left Basem and to-where

- b. šw fkkar-ty min štara šw?
 what thought-2SG.F who bought what
 ‘What did you think? Who bought what?’

D. An embedded wh-phrase can take scope across an indefinite number of wh-scope markers. However, in SA, it is not a requirement that the scope marker be iterated in every clause higher than the embedded wh-phrase. Sentences with only the matrix scope marker, as in (7a) and (78a), are even preferred to sentences with an iterated one, as in (7b) and (7b).

- (7) a. šw btetwak?-i bykun fakkar bassel maʕ min muna kan-et
 what expect-2SG.F be.3SG.M thought Bassel with whom Muna was-3SG.F
 ʕam tħki?
 PROG talking
 Lit. ‘What do you believe? What did Bassel think? Who was Mary talking to?’

- b. ?šw btetwak?-i šw bykun fakkar bassel maʕ min Muna
 what expect-2SG.F what be.3SG.M thought Bassel with whom Muna
 kan-et ʕam tħki?
 was.3SG.F PROG talking

- (8) a. šw fkkar-ty Ali ʔal kif bdna nruħ?
 what thought-2SG.F Ali said how will.1PL go
 ‘How did you think Ali said we will go?’

- b. ?šw fkkar-ty šw ʔal Ali kif bdna nruħ?
 what thought-2SG.F what said Ali how will.1PL go

This has also been pointed out by Beck & Berman (2000). Some speakers of German accept sentences without an iterated wh-scope marker whereas for others an iterated one is preferred. Beck & Berman explain this by assuming that for those speakers who do not require iteration of the wh-scope marker in every clause, the wh-scope marker raises successive cyclically.

E. The complement of the wh-scope marker must be of the interrogative type, whereas the verb must be of the type that requires a [-wh] complement (see Dayal 1994).

As (9) illustrates, a verb which requires a [+WH] complement like *saʔal* ‘ask’ is not acceptable in this construction:

- (9) a. *šw sʔal-ti-a maʕ min kan-et maria ʕam trʔos?
 what asked-2SG.F.SU-3SG.F.OBJ with who was-2SG.F Maria PROG dancing
- b. šw ʔal-ek šw ʕaket Dima?
 what told.3SG.M-2SG what talked Dima

F. Wh-scope marking constructions originating in the complement of factive predicates are grammatical in SA, contrary to the case with full wh-movement questions (compare sentences (a) with sentences (b)):

- (10) a. šw ʕrft min staʔjar l-maʕal?
 what knew.2SG.M who hired the-shop
 Lit. What did you know who will hire the shop?

- b. *min ʕrft staʔjar l-maʕal?
 Who knew.2SG.M hired the-shop

- (11) a. šw ktashaft min shaf Iyad?
 what discovered.2SG.M who saw Iyad
 ‘What did you find out? Who did Iyad see?’

- b. ??Min ktashaf-ty inn-w Iyad shaf?
 Who discovered-2SG.F that-3SG.M Iyad saw

- (12) a. ʕa šw ndm-ti kif ʕkke-ty-a?
 on-what regretted-2SG.F how talked-2SG.F.SU-3SG.F.OBJ

- b. *kif ndm-ti inn-ek ʕkke-ty-a?
 how regretted-2SG.F that-2SG.F talked-2SG.F.SU-3SG.F.OBJ

- (13) a. šw ktashaf-ti kif Sar l-ʕadeth?
 what discovered-2SG.F how happened the-accident
 ‘What did you find out? How did the accident happen?’

- b. *kif ktashf-ty Sar l-ʕadeth?
 how discovered-2SG.F happened the-accident

Contrary to what one may predict for non-referential wh-phrases, of which weak islands obstruct antecedent-government, non-referential wh-phrases originating within the complement of a factive island in scope marking questions are acceptable.

G. Wh-scope marking across negative islands is ungrammatical, as well as full wh-extraction (see sentences (a) and (b)) respectively:

- (14) a. *šw ma fkkart kif raʕ yjw?
 what not thought.2SG.M how will come.3PL

b.*kif ma fkkart inn-w raħ yjw?
 how not thought.2SG.M that.3SG.M will come.3PL

(15)a.*šw ma fkkart maŕ miin Deema kan-et řam thki?
 what not thought.2SG.M with whom Deema was-3SG.F PROG speaking

b. *maŕ miin ma fkkart inn-w Deema kan-et řam thki?
 with who not thought.2sg.m that-3SG.M Deema was-3SG.F PROG speaking

A closer examination shows that some cases of scope marking fail to show negative island effects even with non-referential wh-phrases (see examples (16a, b)). In these examples, scope marking constructions depart from full wh-movement questions.

(16)a. b-řw ma kan yřtref kif fataħ l-bab?
 with-what not was admit.3SG.M how opened the-door
 ‘What did he not admit? How did he open the door?’

b. *kif ma kan yřterf inn-w fataħ l-bab?
 how not was admit.3SG.M that-3SG.M opened.3SG.M the-door

This contradictory behaviour of scope marking constructions originating within negative islands seems to result from the type of matrix predicate, i.e. verbs like *yřtref* ‘admit’ vs. verbs like *yřul* ‘say’, *yřakker* ‘think’. This has also been pointed out by Horvath (1997). A similar behaviour appears in Hungarian scope marking constructions. Horvath argues that the variation with the acceptability is induced by the D-linked vs. non-D-linked status of the propositional complement of the matrix verb. Verbs like *reveal*, *deny*, *notice*, and *permit*, are of the type that takes D-linked interpretation, whereas verbs like *say*, *hear* and *feel*, are of the latter type. The effect of negative islands does not appear with D-linked predicates. It appears with the non-D-linked type of verbs.

To sum up, wh-scope marking in Syrian manifests properties of wh-scope marking. The most significant features of these are: SA allows embedded yes/no questions in wh-scope marking constructions. Wh-scope marking across negative islands is not acceptable with non D-linked predicates; however, it is acceptable with D-linked ones. Wh-scope marking across factive islands is possible, contrary to the case with long extraction questions. Iteration of the scope marker in every clause preceding the embedded clause is not obligatory.

3. Approaches to wh-scope marking

3.1. The direct vs. indirect dependency

The main approaches to scope marking are the *direct dependency* analysis and the *indirect dependency*. The main claim of the direct dependency approach is that scope marking is a variant of extraction structures. Based on facts from German wh-scope marking, McDaniel (1989), following Riemsdijk (1983) argues that the dependency between the wh-scope marker and the embedded wh-phrase is direct. The scope marker is an expletive wh-expression base generated in Spec of matrix CP. It forms a chain with the wh-expression in the embedded CP and its trace, as illustrated in (17):

(17)a. Mit wem_i glaubst du t_i dass Maria t_i gesprochen hat.
 with whom think you that Maria spoken has

- b. Was_i glaubst du mit wem_i Maria t_i gesprochen hat.
 what think you with whom Maria spoken has

(Dayal 1994: 143)

The scope marker is an expletive that is semantically empty. It is base generated in SpecCP in languages that require S-structure wh-movement. Expletives must be eliminated and replaced at LF by a semantically interpreted expression, thus, at LF, the wh-scope marker is replaced by the contentful wh-phrase resulting in an interpretation similar to that of the corresponding long wh-movement question.

According to the indirect dependency approach scope marking and extraction structures are two distinct constructions. They give rise to structural differences, as is the case under negation in German. Whereas full wh-extraction is grammatical across negative islands, as in (18a), wh-scope marking is not possible, as in (18b):

- (18)a. Mit wem glaubst du nicht, dass Maria gesprochen hat?
 With whom think you not that Maria spoken has

- b. *Was glaubst du nicht, mit wem Maria gesprochen hat?
 what think you not with whom Maria spoken has
 ‘Who don't you think Maria has spoken to?’

(Dayal 1994: 145)

Based on facts from Hindi, Dayal argues that the wh-scope marker is generated in argument position. Hindi is an SOV language. The wh-phrase raises from an in-situ position to SpecCP at LF (see examples (19a, b)):

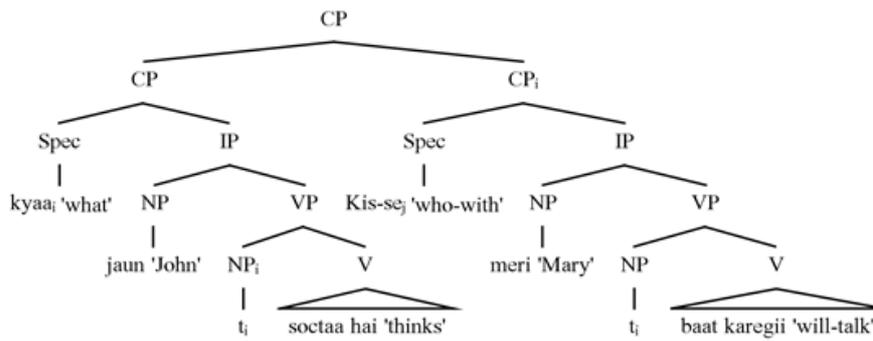
- (19)a. jaun (yeh) jaantaa hai ki meri kis-se baat karegii.
 John this knows that Mary who-with will-talk
 ‘John knows (this) who Mary will talk to.’

- b. Kyaa_i jaunt_i soctaa hai [kis-se_i meri t_j baat karegii]_i;
 what John thinks who-with Mary will-talk
 ‘With who John thinks Mary will talk?’

(Dayal 1994: 150)

In sentences like (19a), the expletive *yeh* ‘this’ is in direct object position, and is coindexed with the complement of the embedded phrase containing the contentful wh-phrase. In analogy with sentences like (19a), Dayal argues that the scope marker in Hindi, *kyaa* ‘what’, in wh-scope marking questions like (19b), originates in object position, and is coindexed with the complement CP. At LF, It moves from the in-situ position to SpecCP, as illustrated in (20):

(20)



3.2. Cross linguistic variation

Based on facts from *wh*-scope marking in Hungarian, Horvath (1997, 2000) argues that in languages like Hungarian, the scope-marker is an expletive base-generated in a non-theta A-position. The embedded CP originates in the argument position of the matrix verb. It raises at LF to adjoin to the expletive *wh*-scope marker in matrix CP, as in (21b).

- (21) a. Mit gondolsz, hogy kit látott János?
 what-ACC think-2SG that who-ACC saw-3SG John-NOM
 ‘Who do you think that John saw?’

(Horvath 1997: 510)

- b. [CP [CP hogy kit látott János] mit [C gondolsz t_{CP}]]
 that who-ACC saw-3SG John-NOM what-ACC think-2SG

This can be supported by evidence from the case system in Hungarian. The case attested on the scope marker is determined by the grammatical relation that the embedded clause has with respect to the predicate in the matrix clause. The *wh*-scope marker bears an accusative case only when the complement clause is a direct object argument of the matrix predicate, as illustrated in (22). The agreement appearing in the intermediate clause is an indication that *mit* ‘what’ must have been in that clause before it has moved to the matrix clause.

- (22) Mit_i mondt_ā [hogyt_i tudnak/ *tudj_ā,
 what.ACC said.2SG.INDEF.DO that know.3PL.INDEF.DO/ *know.3PL.DEF.DO
 hogy melyik fiut szereted t_i]]?
 that like.2SG.DEF.DO which boy.ACC
 ‘Which boy did you say that they know that you like?’

Horvath argues that languages manifest different properties in terms of *wh*-scope marking cross-linguistically, thus they require different analyses. However, Dayal (2000) argues that scope marking is a universal phenomenon. Assuming that languages differ with their syntactic rather than semantic realizations of *wh*-scope marking, the different characteristics can be reconciled under an indirect dependency analysis with a variety of approaches.

3. Approaches to wh-scope marking in Syrian Arabic

3.3. An indirect dependency

Properties of wh-scope marking in SA suggest that there is no direct dependency between the wh-scope marker and the embedded wh-phrase. The fact that a yes/no question can be embedded in a wh-scope marking question shows that the wh-scope marker and the embedded wh-phrase do not form a wh-chain, as illustrated in sentence (5). Another piece of evidence that supports this result is the fact that embedded questions with more than one wh-expression can occur in the embedded clause, as in (6). It is argued in Dayal (1994) that these questions raise a problem for the direct dependency between the scope marker and the embedded wh-phrase. The wh-chain would have one head and two tails. However, this is not a problem for the indirect dependency approach. Each wh-expression can be interpreted at its LF position.

The other argument against the direct dependency analysis follows from the contrastive behaviour of wh-scope marking constructions and extraction questions involving factive predicates, as in (10-13), and negative islands with D-linked verbs, as in (16). In these examples, wh-scope marking questions depart from full wh-movement questions.

From what has been discussed, it can be concluded that wh-scope marking in SA is not isomorphic to extraction structures, and the dependency between the wh-scope marker and the embedded wh-clause is indirect.

3.4. Split constituents

Another analysis of wh-scope marking suggests that the wh-scope marker and the embedded question form a DP constituent in the underlying syntactic structure but split in the course of the derivation, the wh-scope marker undergoing wh-movement to SpecCP. This would be similar to the *was-für* split construction in German illustrated in (23) (see Herburger 1994, Bruening 2004, Leu 2008):

(23) a. [Was für ein Buch] hast du gelesen?
What for a book have you read
'What kind of/which book did you read?'

b. [Was] hast du [für ein Buch] gelesen?
what have you for a book read

The wh-scope marker would be the head D, whose sister is the embedded CP, which provides restriction for the wh-scope marker, as illustrated in (24):

(24) [CP what_i [TP you [VP say [DP t_i [CP who_j [you saw t_j]]]]]]

The head would separate from the restricting CP and move to the matrix CP. The wh-word and the CP would originate in the same position, as one complement to the matrix verb, which becomes a split constituent due to movement of its head, the wh-scope marker. It cannot actually be movement of a head in the X-bar sense, though, since the movement does not have the properties of head-movement but of A-bar movement. Instead it would be movement of a 'maximal wh-word', a counterpart of *what*, which functions as a determiner of a clausal complement¹.

¹ Bruening (2004) argues against this objection (see Bruening 2004: 284).

Whatever the best analysis is of *was-für* split, this seems to be a rare phenomenon. There are cases reported in the literature of determiners or quantifiers splitting from their complement, but the most common ones move the complement stranding the determiner/quantifier. This is the case with quantifier stranding (or quantifier float) (Sportiche 1988).

In SA, split DP constituents are not attested at all. A determiner or a quantifier cannot appear separately from its NP restriction. There is nothing corresponding to *was-für* split. There is not even quantifier stranding. One kind of split constituent that we do see, in many languages is with extraposed relatives. (25a) derives from the underlying structure (25b).

- (25) a. Everything is true that she told me.
 b. [IP is true [DP everything that she told me]]

Even this split is not possible in SA (see (26)):

- (26) a. kl lli ħaket-w mazbut.
 every that said.3SG.F-it true
 ‘Everything that she said is true’
- b. mazbut kl illi ħaket-w.
 true every that said.3SG.F-it
 Intended: true everything that she said.
- c. *kl mazbut lli ħaket-w.
 every true that said.3SG.F-it
 Intended: everything is true that she said.

Note that it is possible to move the relative clause along with the head, as in (26b) in SA. This suggests that, if it were true that the wh-scope marking construction is derived by splitting a DP consisting of a determiner and a restrictive clause, we should expect movement of the entire DP, as an option. This is not possible in SA, or in any language with wh-scope marking, as far as I know (see (27)):

- (27) *šū maʕ miin akl-et ħaneen btzn-i?
 what with who ate-3SG.F Haneen think-2SG.F
 Intended: *What with whom ate Haneen you think?

As can be seen, there is little reason to think that wh-scope marking in SA arises from a split DP constituent. Instead, the wh-scope marker and the embedded CP are distinct constituents.

3.5. The predication analysis

3.5.1. The wh-scope marker as an argument

According to the standard analysis of wh-scope marking, the wh-scope marker is an expletive base generated in Spec of matrix CP (Riemsdijk 1983, McDaniel 1989). However, it is argued in Dayal (1994) that the wh-scope marker is an expletive base-generated in matrix object position. The wh-scope marker *was* in German is the [+wh] counterpart of the CP expletive *es* ‘it’ in structures like (28). This assumption is supported by the observation that *was* never co-occurs with [-wh] CP-expletives.

(28)a. **Es** wird behauptet [CP dass Maria Mel Gibson getroffen hat]
 it is said that Maria Mel Gibson met has
 ‘They are saying that Maria has met Mel Gibson.’

b. **Was** wird (***es**) behauptet [CP **wen** Maria getroffen hat]?
 what is it said whom Maria met has
 ‘Who are they saying Maria has met?’

In line with Rothstein’s (1995) claim that true object expletives do not exist, expletives can only originate in subject positions since they get licensed through syntactic predication rather than θ -marking. Felser (2001) argues that the *wh*-scope marker *was* is not an expletive but an internal argument of the matrix verb. It is excluded from positions reserved for true expletives. It cannot substitute for the expletive *es* ‘it’ in sentences like (29):

(29)a. Er findet [SC es *überraschend*] [dass Maria Hans noch liebt]
 he considers it surprising that Maria Hans still loves
 ‘He considers it surprising that Maria still loves Hans.’

b. ***Was**_i findet er [SC t_i *überraschend*] [**wen**_k Maria t_k noch liebt]?
 what considers he surprising whom Maria still loves

A similar point can be made about the *wh*-scope marker in SA, even though the facts are a bit more complicated. SA has a counterpart of the German expletive, but constructed with a complementizer, as shown in (30):

(30) ilt-l-ek inn-w_i [Deena ma rah tj-i l-yom]_i
 told.1SG.SU-to-2SG.F.OBJ that-3SG.M Deena not will come-3SG.F the-today
 ‘I told you that Deena is not going to come today.’

The complementizer *inn-* also co-occurs with referential pronominal clitics, as in (31).

(31)a. baʔref inn-ek zʔlaneh.
 know.1SG that-you upset
 ‘I know that you are upset.’

b. baʔref inn-a ma rah tji.
 know.1SG that-her not will come
 ‘I know that she is not going to come.’

This complementizer is employed in Standard Arabic in the form of *ʔanna*, as in (32). It introduces both finite and non-finite clauses, and assigns accusative case to the following noun or pronoun (See Aoun et al 2010: 17; Mohammad 2000: 108). When not immediately followed by a lexical subject, as is the case in (32), it is followed by an expletive, an accusative clitic bearing default masculine singular form.

(32)a. Qultu ʔinna-hw wasala l-ʔ-wlaad-u [Standard Arabic]
 said.1SG that-it arrived.3SG.M the-children-NOM
 ‘I said that the boys arrived.’

(30) Has quite a different structure, though. The clitic element *-w* does not substitute for a subject, or take the subject as associate; the embedded clause has an initial lexical subject. Instead, I assume, the clitic is a clausal expletive, in that sense similar to *es* in (28a). It is coindexed with the clause *Deena ma raħ tji lyom* ‘Deena will not come today’. The underlying structure is [DP-*w* [CP inn-IP]], and the surface morpheme order is derived by incorporation of the complementizer in the nominal head *-w*.

The wh-scope marker *šu* can still co-occur with the expletive in sentences like (33). This shows that the wh-scope marker does not originate in the same position of the expletive.

- (33) *Šu fkkart ʔnn-w miin maria bthb?*
 what thought.2SG.M that-3SG.M who Maria love
 ‘Who did you think that Maria loves?’

This result does not entail that the wh-scope marker is base generated in Spec of the matrix CP. *Šu* is a propositional proform as in *šu fkkart?* ‘What did you think?’. It can be the direct object of the matrix verb and is capable of carrying the role of Theme. This, I claim, is the case in (33).

Another piece of evidence is that the wh-scope marking strategy is unavailable when the complement of V is a complex DP, as in (34),

- (34) **šu fkkart fkra maħ miin Deema kan-et ʕam thk-i?*
 what thought.2SG.M idea with whom Deema was-3SG.F PROG speaking
 Intended: ‘Who did you have an idea that Deema was talking to?’

This can be understood if *šu* ‘what’ in the scope-marking construction is an argument of the verb base generated and receiving a theta-role in the direct object position. If so, (34) violates the theta-criterion.

Following from this discussion, it can be concluded that the wh-scope marker is not an expletive base generated in Spec of the matrix CP, nor an expletive in construction with an embedded clause, but an argument of the matrix verb.

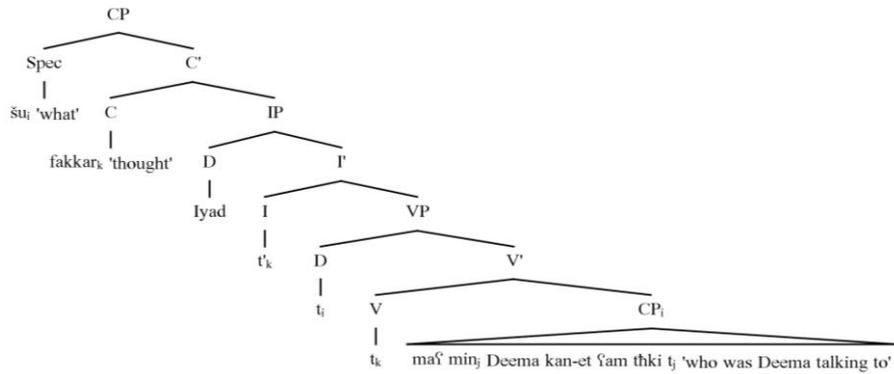
3.6. Secondary predication

Following Felser’s (2001), I propose that wh-scope marking constructions in SA can best be understood in terms of a predication relation between the wh-scope marker and the embedded wh-clause, referred to by Felser (2001) as a case of secondary predication. The wh-scope marker is an object pronoun that originates in Spec of VP where it is assigned the θ -role of Theme. It moves to Spec of matrix CP to check its interrogative feature.

The embedded wh-clause, therefore, is not assigned a thematic role by the verb, but is an ‘unselected complement’ of the verb. This is the basis for the relation between the wh-scope marker and the wh-clause.

- (35) *šu_i fakkar Iyad [maħ min Deema kan-et ʕam thki]_i?*
 what thought Iyad with whom Deema was-3SG.FPROG speaking
 ‘What did Iyad think? Who was Deema talking to?’

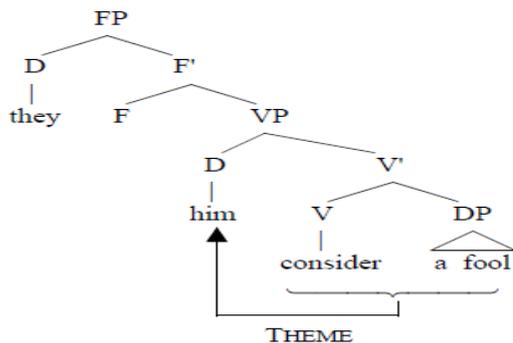
(36)



The wh-scope marker *šw* ‘what’ originates in Spec of VP where it is assigned the θ -role of Theme. It raises to Spec of matrix CP. The verb *fakkar* ‘thought’ takes the CP *maʕ min Deema kanet ʕam thki* ‘with whom Deema was speaking’ as its unselected complement (see Felser 2001).

Felser (2001) likens this analysis to the relation between *him* and a *fool* in the small clause construction *They consider him a fool*, but analysed as in Williams (1997) as made up, underlyingly, of a complex predicate [consider a fool] assigning an object role to him.

(37)



In this structure, *a fool* would be an unselected complement of *consider*, and the resulting interpretation is that *him* and *a fool* enter a predication relation.

According to this analysis, there is no direct dependency between the scope marker and the wh-phrase in the embedded clause. There is no LF replacement of the wh-scope marker by the embedded CP; rather they are in subject-predicate relation. If the wh-scope marker undergoes wh-movement to Spec of CP, this explains the fact that wh-scope marking is incompatible with islands. And contrary to the case in Hungarian, there would be no movement at LF in SA, where the embedded CP moves along with the island to adjoin the wh-scope marker causing no violation of movement out of the island.

This analysis is also compatible with the fact that wh-scope marking constructions in SA allow embedded yes/no questions. Since the wh-scope marker and the embedded clause form separate chains (they are related by predication, not by movement), an embedded yes/no question does not lead to any violation.

The fact that the embedded clause is a question seems to satisfy the requirement that the associate of the wh-scope marker has to be a [+WH] interrogative, as is argued in Dayal (1994). The acceptability of an embedded yes/no question shows that the [+WH] associate has to be an interrogative but not a wh-question necessarily.

As regards clauses without copies of the wh-scope marker in multiple embedded clauses, in such sentences, the wh-scope marker undergoes long wh-movement from the intermediate clause to Spec of the matrix clause.

While it is not intuitively obvious that the relation between *šw* ‘what’ and the *wh*-clause in (35) is the same as the relation between *him* and a *fool* in (37), it does seem that we can regard it as a form of identificational relation (Higgins 1973), comparable to the relation between the two terms in (38):

(38) That person is our leader.

The wh-clause provides identification or definition of the otherwise completely underspecified wh-pronoun *šw*. I will discuss this further in a future research.

5. Conclusion

In this paper, I have argued that Syrian Arabic makes use of the wh-scope marking strategy as an alternative to long movement questions for questioning out of embedded questions. I argued that the dependency between the wh-scope marker and the embedded wh-phrase is indirect. The wh-scope marker and the embedded clause do not form a constituent at either the underlying structure or at LF. The wh-scope marker originates in the direct object position of the matrix verb and raises to Spec of matrix CP. Following Felser (2000), I have argued that the embedded clause is an unselected predicate of the verb.

The properties of wh-scope marking in SA are best interpreted in terms of an identificational relation, in which the embedded clause provides identification of the wh-scope marker. This analysis builds into Dayal’s (2000) conclusion that the different varieties of wh-scope marking constructions cross-linguistically can be reconciled under a unified approach, the indirect dependency.

References

- Aoun, Joseph E., Elabbas Benmamoun, and Lina Choueiri. (2009) *The syntax of Arabic*. Cambridge University Press.
- Bailey, L., Holmberg, A., Krzek, M., Sheehan, M. and Sulaiman, M. (2010) Intonation questions. *The 2010 Annual Meeting of the Linguistics Association of Great Britain*. Leeds, 1st–4th September (2010).
- Beck, S. and Berman, S. (2000) Wh-scope marking: Direct vs. indirect dependency. *Wh-scope marking*, 17-44.
- Bruening, B. (2004) Two types of wh-scope marking in Passamaquoddy. *Natural Language & Linguistic Theory*, 22(2), 229-305.
- Dayal, V. (2000) Scope marking: Cross-linguistic variation in indirect dependency. *Wh-scope marking* 37: 157-193.
- Dayal, V.S. (1994) Scope marking as indirect wh-dependency. *Natural language semantics*, 2(2): 137-170.
- Felser, C. (2001) Wh-expletives and secondary predication: German partial wh-movement reconsidered. *Journal of Germanic Linguistics*, 13(01), 5-38.
- Herburger, E. (1994) A Semantic Difference between Full and Partial Wh-movement in German. *Annual Meeting of the Linguistic Society of America*. Boston.
- Higgins, F.R. (1973) *The pseudo-cleft construction in English*. Doctoral thesis, Massachusetts Institute of Technology.

- Horvath, J. (1997) The status of 'wh-expletives' and the partial wh-movement construction of Hungarian. *Natural Language & Linguistic Theory*, 15(3): 509-572.
- Horvath, J. (2000). On the syntax of 'Wh-Scope Marker' Constructions: Some Comparative Evidence. *Wh-scope marking*, 37:271.
- Lahiri, U. (2002) On the proper treatment of “expletive< i> wh</i>” in Hindi. *Lingua*, 112(7), 501-540.
- Leu, T. (2008) What for internally. *Syntax*, 11(1), 1-25.
- McDaniel, D. (1989) Partial and multiple wh-movement. *Natural Language & Linguistic Theory*. 7(4), 565-604.
- Mohammad, M. (2000) *Word order, agreement, and pronominalization in Standard and Palestinian Arabic*. Vol. 181. John Benjamins Publishing,
- Riemsdijk, H.v. (1983) Correspondence effects and the empty category principle, in Otsu et al (eds.), *Studies in Generative Grammar and Language Acquisition*, Tokyo: International Christian University, 5-16.
- Rothstein, S. (1995) Pleonastics and the interpretation of pronouns. *Linguistic Inquiry*, 26, 499–529.
- Williams, E. (1997) Lexical and syntactic complex predicates. *Complex predicates*, (64), 13-28.

*Mais Sulaiman
Department of Linguistics
Newcastle University
Newcastle Upon Tyne
United Kingdom
NE1 7RU*

m.sleman1503@gmail.com

THE INTERVENTION EFFECT OF NEGATION ON WH-ADVERBIALS IN LATE ARCHAIC CHINESE¹

AIQING WANG
(University of York)

Abstract

This paper investigates the Intervention Effect of negation in Late Archaic Chinese (5th-3rd c BC; ‘LAC’). *Wh*-complements of adverbials must raise out of the head-initial PPs to a position between TP and *v*P due to obligatory *wh*-preposing, generating the reverted order *wh*-P. Fronted *wh*-constituents may target one of the two focalised positions in the medial domain, and the distributional asymmetry of *wh*-phrases is correlated with their base positions. The lower focused position below negation accommodates preposed *wh*-adverbials base-generated between negation and *v*P, while the higher focus position above negation is expected to exclusively permit *wh*-constituents base-generated above negation, viz. *wh*-complements of high PPs. However, when a negator is present and c-commands a *wh*-adverbial that is supposed to target the low focus position, it will trigger *wh*-movement to the high focused position due to the Intervention Effect of negation (Beck 1996, Beck and Kim 1997, Kim 2002, 2006).

1. Introduction

The blocking effect in the sense of Beck (1996) and Beck and Kim (1997) refers to the fact that a barrier may not intervene between a question existential operator (Q-operator) and a function variable bound by that Q-operator. Such a blocking effect applies to *wh*-in-situ, the stranded restriction of *wh*-constituents moving overtly, as well as *wh*-scope marking structures. Kim (2002, 2006) analyses the blocking effect on LF (covert) dependency of *wh*-in-situ as a function variable bound by a Q-operator (termed by Hagstrom (1998) and Pesetsky (1999) as the Intervention Effect), and proposes that it is a focus phrase that induces an Intervention Effect in modern Mandarin, instead of negation or quantifiers in general. Additionally, there is a repair strategy to circumvent the Intervention Effect in modern Mandarin by means of raising *wh*-in-situ to a position preceding the focus-induced barrier, as exemplified by (1).

- (1) a. ?**Zhiyou* Lili kan-le na-ben shu? (Kim 2002: 626)
only Lili read-ASP which-CL book

* I would like to express my sincere gratitude to my supervisor, Professor Peter Sells, for his inspiring guidance and instruction on this paper. I am especially indebted to Dr Shin-Sook Kim for her invaluable and helpful comments and feedback on the original draft.

- b. Na-ben shu *zhiyou* Lili kan-le?
 which-CL book only Lili read-ASP
 ‘Which book did only Lili read?’

Contrary to this, quantifiers or negation do not show the Intervention Effect. Consequently, a quantified structure or a negator can take a position between a Q-operator and a *wh*-in-situ bound by this Q-operator, as in (2a/b/c) that contain an ordinary quantifier NP, a frequency adverbial and a negator respectively (Kim 2002).

- (2) a. Meigeren dou mai-le shenme? (Kim 2002: 625)
 everyone all buy-ASP what
 ‘What did everyone buy?’
- b. Zhangsan changchang mai shenme?
 Zhangsan often buy what
 ‘What does Zhangsan often buy?’
- c. Zhangsan bu xiang mai shenme?
 Zhangsan not want buy what
 ‘What doesn’t Zhangsan want to buy?’

In this paper I investigate the Intervention Effect of negation in Late Archaic Chinese (LAC) that refers to Archaic Chinese during the Warring States period (5th-3rdc BC).² Chinese has always been an SVO language from the earliest textual sources viz. the Shang inscriptions (14th-11thc BC) to modern Mandarin (Peyraube 1996, Djamouri 2001, Djamouri and Paul 2009, Meisterernst 2010, Djamouri et al 2012), and texts in LAC display predominant SVO word order. However, there are contexts in which nominal and pronominal object DPs appear preverbally in the ‘low IP area’ (dubbed by Paul (2005)), as shown in (3a-b),³ which involve nominal and pronominal non-*wh*-objects respectively. 之 ZHI in (3) functions as an optional fronting marker accompanying preposed DPs. Various observations support the view that object preposing in LAC is derived, and should not be presumed as the vestige of basic OV word order.

- (3) a. 吾 百姓 之 不 圖 (Guoyu)
 wu baixing_i zhi bu [_{VP} tu t_i]
 I common.people ZHI not care.about
 ‘I did not care about common people’

² I follow Wang (1958), Zhou (1963) and Peyraube (2003) in terming Classical Chinese during the Warring States period as Late Archaic Chinese (LAC), which exhibits distinctive characteristics. I also agree that around the Han Dynasty (2ndc BC-2ndc AD) after the Warring States period, there was a transitional period with multiple typological changes (Xu 2006).

³ The primary sources of LAC examples in this paper are Peking University corpus, Academia Sinica electronic database, and the Sheffield Corpus of Chinese. The selected texts of these corpora are all received, representing a wide range of writing found in various time periods. In LAC period, the corpora cover more than twenty key books written by different authors.

- b. 吾 斯 之 未 能 信。 (Analects)
 Wu si_i zhi wei neng [_{VP} xin t_i].
 I this ZHI not.yet can be.confident
 ‘I have not been able to be confident in this yet.’

Furthermore, if the objects or other VP-internal constituents are *wh*-DPs, they normally front to a position below the subject and above *v*P, because LAC was a *wh*-fronting language, and *wh*-in-situ did not emerge in Chinese until the Han Dynasty (2ndc BC-2ndc AD) (Feng 1996, Aldridge 2010). Examples (4a) and (4b) show *wh*-DPs functioning as a direct object and a locative element respectively, both of which raise to a preverbal position in LAC.

- (4) a. 宋 何 罪 之 有? (Mozi)
 Song [he zui]_i zhi [_{VP} you t_i]?
 Song what in ZHI have
 ‘What sin does the State of Song have?’

- b. 其 子 焉 往? (Mencius; Aldridge 2013: 246)
 Qi zi yan_i [_{VP} wang ti]?
 3.GEN son where go
 ‘Where will their sons go?’

I analyse *wh*-fronting in LAC, and propose that the Intervention Effect of negation is present in LAC. This paper is organised into two main sections. In Section 2 I introduce *wh*-fronting in LAC, including *wh*-complement of reason PPs and other adjunct adverbials. In Section 3 I examine the Intervention Effect on *wh*-adverbials and propose that negation exhibits such a blocking effect.

2. *Wh*-Fronting

Since LAC was a *wh*-fronting language, *wh*-complements of adverbials undergo obligatory movement and front out of the head-initial PP, generating the derived *wh*-P order.⁴ I explore the preposing of *wh*-adverbials in LAC, and propose that *wh*-phrases front to focalised positions in the ‘low IP area’. *Wh*-complements of ‘high’ adverbial PPs ‘why’ raise to a high focus above negation, yet *wh*-complements of other adjunct adverbials target a low focus position intervening between negation and *v*P.

2.1. *Wh*-Complements of Reason Adverbials

Wh-complements of reason adverbials undergo obligatory movement. In (5), a simplex *wh*-complement 何 *he* fronts out of a head-initial PP 為何 *wei he* ‘for what’ that functions as an adverbial of reason above negation, generating the inverted *wh*-P order.

⁴ I argue that such inverse ordering is generated via separate movement of *wh* and the preposition, rather than inversion with PP, mere *wh*-fronting or pied-piping. See Wang (2013) for detailed discussion.

(5) a. 吾 何 為 不 增? (Guoyu)
 Wu he_j wei_i [PP t_i t_j] bu [VP zeng]?
 I what for not enhance
 ‘For what do I not enhance (it)?’

b. 吾 何 為 獨 不 然? (Mencius)
 Wu he_j wei_i [PP t_i t_j] du bu ran?
 I what for alone not correct
 ‘For what am I alone not correct?’

The base position of ‘high’ reason adverbials in LAC is above negation, hence above *vP*, and the reversed order *wh*-P indicates that movement has happened. Therefore, I suggest that *wh*-complements of reason adverbials are base-generated above negation and move to an even higher position. In LAC, non-D-linked *wh*-constituents consistently display focus properties, so the landing site for adverbials of reason should be a focalised position, referred to as the High focus position.

Given the fact that reason adverbials are the only ‘high’ adverbials in LAC, the High focus position might be expected to allow reason *wh*-adverbials exclusively.

2.2. *Wh*-Complement of Other Adjunct Adverbials

A *wh*-complement of an adjunct PP also fronts to a preverbal position, triggered by obligatory *wh*-preposing. Some adjunct adverbials can be base-generated in a position between negation and *vP*, yet others are base-generated in a postverbal position. This observation coincides with the generalisation from Djamouri et al. (2012) that adjunct PPs in Classical Chinese may occur in a pre- or postverbal position. *Wh*-adverbials base-generated between negation and *vP* or postverbally always undergo movement and land in a focused position intervening between negation and *vP*.

First, the *wh*-complement of an adverbial of instrument undergoes obligatory *wh*-movement. In (6a), the same *wh*-complement 何 *he* as that in (5) raises out of a head-initial adjunct PP 以何 *yi he* ‘with what’ to the medial domain between TP and *vP*. Again, the reversed *wh*-P order proves that movement has occurred. As can be seen from a canonical example involving a non-*wh*-PP, instrumental adverbials are base-generated lower than negation and higher than *vP* (6b).⁵

(6) a. 將 何 以 守 國? (Guoyu)
 Jiang he_j yi_i [PP t_i t_j] [VP shou guo]?
 FUT what with guard state
 ‘With what will (he) guard the state?’

⁵ In addition to being a fronting marker (see (3) and (4a)), the same graph 之 *zhi* can also act as a third person accusative pronoun.

- b. 君王 不 以 鞭笞 使 之。 (ibid)
 junwang bu [PP yi bianchui] [VP shi zhi]
 Your.Majesty not with whip enslave 3.OBJ
 ‘Your Majesty did not enslave them with a whip’

Second, a manner adverbial is also base-generated preverbally, with its *wh*-complement undergoing obligatory movement. Analogous to instrumental adverbials, moved manner adverbials can also be diagnosed by the inverted PP 何以 *he yi* (7a). (7b) shows a canonical example involving the basic order Neg-PP-VP.

- (7) a. 子 何 以 知 之? (Zuozhuan)
 Zi he_j yi_i [PP t_i t_j] [VP zhi zhi]?
 you what with know 3.OBJ
 ‘How do you know it?’

- b. 不 以 其 道 得 之 (Analects)
 bu [PP yi [qi dao]] [VP de zhi]
 not with 3.GEN means obtain 3.OBJ
 ‘obtain them not by their means’

Third, the base position of source adverbials can be postverbal, and the *wh*-element moves from the postverbal position to a preverbal position triggered by obligatory *wh*-preposing, as in (8a). I assume that the source adverbial in (8a) involves an empty preposition 于 *yu* ‘from’.⁶ A supporting example for this view is (8b) where a question and its answer demonstrate structural similarity. As can be seen from the answer, source information is situated after the verb. So it is safe to assume that in the default, unmarked order, the *wh*-word 惡 *wu* ‘where’ concerning the source information in the question should also appear after the verb.

- (8) a. 仲尼 焉 學? (Analects)
 Zhongni yan_i [VP xue [PP ti]]?
 Zhongni where study
 ‘(From) where does Zhongni study?’

⁶ The unmarked counterpart of (8a) is in (i) where the source PP headed by 于 *yu* is base-generated postverbally. This instance coincides with Peyraube’s (1997) generalisation that 于 *yu* is a preposition with a relatively rigid postverbal position.

(i) 且 子 獨 不 聞 夫 壽陵 餘子 之 學 于 邯鄲 歟?
 Qie zi du bu wen fu Shouling yuzi zhi [VP xue [PP yu Handan]] yu?
 CONJ you alone not hear DET Shouling youth ZHI learn from Handan Q
 ‘Besides, did you alone not hear of that youth from Shouling who learned (the manner of walking) from Handan?’ (Zhuangzi)

- b. 惡 乎 取 之? 取 之 曹 也。
 Wu_j hu_i [VP qu zhi [PP t_i t_i]]? [VP Qu zhi [PP Cao]]ye.
 where from take 3.OBJ take 3.OBJ Cao DECL
 ‘From where to take it? Take it (from) Cao.’

(Gongyangzhuan)

Fourth, when locative *wh*-adverbials are right-adjoined to VP, they front to a preverbal position triggered by obligatory *wh*-fronting. As suggested by Peyraube (1996), locative PPs are predominantly postverbal in Archaic Chinese. Example (9) contains a question and answer pair. Since in the answer the locative PP (also involving a null preposition) follows the VP, it is natural to posit that the *wh*-PP in the question is also base-generated postverbally, and the surface order *wh*-P-VP is derived from *wh*-movement and P movement.

- (9) 惡 乎 用 之? 用 之 社 也。
 Wu_j hu_i [VP yong zhi [PP t_i t_i]]? [VP Yong zhi [PP she]] ye.
 where in use 3.OBJ use 3.OBJ shrine DECL
 ‘In where (did he) use him? (He) used him (in) the shrine.’

(Gongyangzhuan)

Before demonstrating the landing site of *wh*-complements of adjunct adverbials other than reason PPs, I need to show that there is a key adverb 獨 *du* ‘alone’ that can diagnose the position where negation is generated. My observation reveals that the adverb 獨 *du* ‘alone’ always immediately precedes negation, and no element can intervene between *du* and the following negative, as in (10) (also refer to (5b) and (i)). So if a *wh*-element precedes or follows *du*, then this *wh* must precede or follow the position of negation accordingly. Therefore, *du* is a crucial diagnostic element to decide the relative order between *wh*-phrases and negation, even without the presence of a negator.

- (10) 子 獨 不 聞 涸 澤 之 蛇 乎? (Hanfeizi)
 Zi du bu [VP wen [he ze zhi she]] hu?
 You alone not hear.of dry marsh GEN snake Q
 ‘Have you alone not heard of (the parable about) snakes in a dry marsh?’

When the *wh*-complement of an adjunct PP raises to a higher position triggered by obligatory *wh*-preposing, it is still below negation, because it follows the adverb 獨 *du* that always immediately precedes negation (11). So despite of the absence of negation in instances involving adjunct adverbials other than reason PPs, the *wh*-complements are expected to target a position below negation and above *v*P. Due to the non-D-linked nature of these *wh*-adverbials, the landing site should be a focus position, which I call the Low focus position. (11a) and (11b) illustrate manner and source PPs respectively.

- (11)a. 先生 獨 何 以 說 吾 君 乎?
 Xiansheng du he_j yi_i [PP t_i t_j] [VP yue wu jun] hu?
 sir(you) alone what with please my lord Q
 ‘How were you alone able to please my lord?’

(Zhuangzi; Aldridge 2010:10)

- b. 子 獨 惡 乎 聞 之?
 Zi du wu_j hu_i [VP wen zhi [PP t_i t_j]]?
 you alone whom from hear 3.OBJ
 ‘From whom did you alone hear it?’

(Zhuangzi)

To summarise, the High focus position above negation allows ‘high’ *wh*-adverbials exclusively, and the Low focus position below negation accommodates other adjunct *wh*-adverbials, including instrument and manner PPs base-generated between negation and *vP*, as well as source and locative PPs base-generated postverbally. So I conclude that for the preposing of *wh*-adverbials, their landing sites are correlated with and determined by their base positions, shown in the template as follows:

- (12) Clausal positions of *wh*-fronting:

Subject > High focus position > High *wh* base position > Negation > *du* > Low focus position > Low *wh* base position > *vP*

3. The Intervention Effect of Negation

In this section, I investigate the presence of the Intervention Effect of negation on focalised *wh*-adverbials in the Low focus position.

3.1. Preposed *Wh*-Adverbials

Wh-adverbials other than reason PPs are base-generated preverbally or postverbally, and they always front to the Low focus position. So far, examples involving the Low focus position contain no negatives, and we do not expect negation to make any difference based on data in modern Mandarin. Nevertheless, such a prediction is not borne out for LAC, in that if a negative element is present and *c*-commands an interrogative *wh*-phrase, the negator will block the LF dependency of the *wh*-constituent, due to the Intervention Effect (Beck 1996; Beck and Kim 1997; Kim 2002, 2006). To be in a position where it can be interpreted semantically, the *wh*-element has to adopt a repair strategy by fronting to a position across negation, until the blocking effect is circumvented. As a consequence, the surface landing position of a *wh*-constituent is always above negation, corresponding to Aldridge’s (2010) generalisation that *wh*-words never follow negators.

Preposed *wh*-adverbials in the Low focus position below negation have to raise again to an operator position as their LF dependency cannot cross a negative barrier. Both source and locative PPs *wh*-complements are affected by the Intervention Effect of negation.

Wh-phrases functioning as source adverbials are subject to the Intervention Effect of negation. As illustrated earlier, source adverbials can be base-generated postverbally and their *wh*-complements move to a position between negation and *vP*, following the key diagnostic element 獨 *du*, as in (13a) (=11b)). However, in the context of negation, the *wh*-adverb 焉 *yan* ‘where’ in (13b) moves overtly from its postverbal base position to the High focus position across the negative. It can be seen that negation functions as a barrier for the interpretation (Q-binding) of *wh*-adverbials base-generated above *vP*: they would target the Low focus position if there was no Intervention Effect of negation. However, the Low focus position cannot accommodate these *wh*-PPs, as it is c-commanded by the negator, parallel to the postverbal base position. So instead of raising to the Low focus position and still being c-commanded by the intervener, *wh*-elements target the High focus position c-commanding the negative intervener.

(13) a. 子 獨 惡 乎 聞 之? (Zhuangzi)
 Zi du wu_j hu_i [VP wen zhi [PP t_i t_j]]?
 You alone whom from hear 3.OBJ
 ‘From whom did you alone hear it?’

b. 夫子 焉 不 學? (Zuozhuan)
 Fuzi yan_i bu [VP xue [PP ti]]?
 Confucius where not study
 ‘(From) where does Confucius not study?’

Similarly, locative PPs are base-generated postverbally and their *wh*-complements are expected to target a landing site between negation and *vP*. However, with the presence of negation, location PPs always appear in a position preceding the negator (14). This fact also lends support to the proposal that the intervening negator blocks LF movement of a *wh*-element to an operator position, so the *wh*-element has to raise to a position c-commanding the negative barrier.

(14) 子 何 所 不 逞 欲? (Zuozhuan)
 Zi [he suo]_i bu [VP cheng yu [pp ti]]?
 you what place not satisfy desire
 ‘(In) what place do you not satisfy desires?’

It is notable that the High focus position not only accommodates *wh*-words as ‘high’ adverbials, but other non-D-linked *wh*-adverbials to move to this position as well, due to the Intervention Effect of negation, as illustrated by examples (13b) and (14).

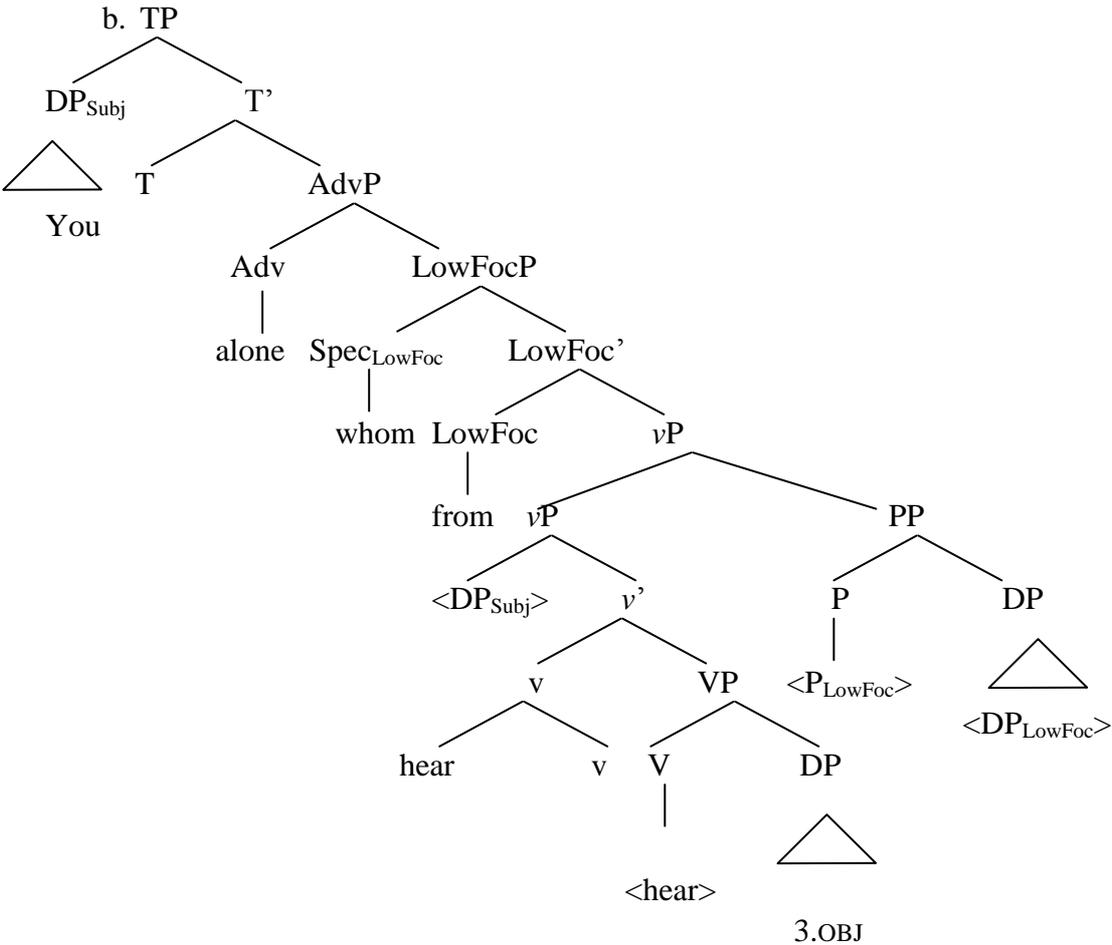
3.2. Landing Sites of *Wh*-Adverbials

In terms of the landing sites of *wh*-fronting, I adopt the theory that a preposed *wh*-adverbial occupies a specifier position of some functional category below TP and above *vP* (Paul 2002, 2005). Owing to the focalised nature of two positions for *wh*-preposing, they are

termed as HighFocP and LowFocP respectively; and the optional fronting marker ZHI accompanying *wh*-DPs is thus a focus marker. Both the focus marker ZHI and prepositions which move out of PP target the head of these functional projections, so that is why there is complementary distribution of the focus marker ZHI and prepositions. To be more specific, the reason why the focus marker only follows a focused DP (see (3) and (4a)) but never coexists with a focused PP is because when the head node of a focus phrase (either the HighFocP or LowFocP) is occupied by a fronted preposition (see e.g. (5), (9) and (11))), there is no position for a focus marker, and vice versa (4a).

The tree diagram of (15a) (=13a) without the Intervention Effect of negation is (15b), where the fronted *wh*-complement 惡 *wu* ‘whom’ occupies the specifier of the functional projection LowFocP, and the preposition 乎 *hu* ‘from’ moves to the head position.

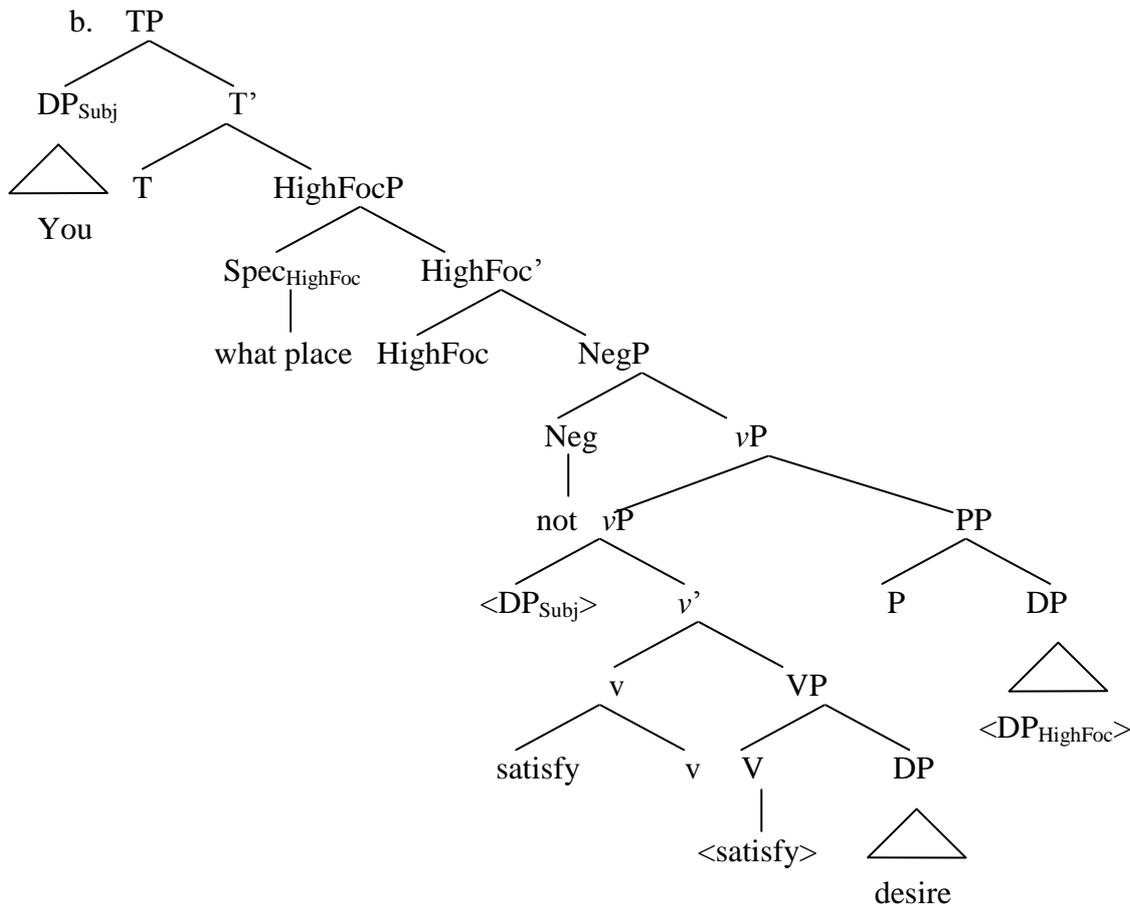
(15)a. 子 獨 惡 乎 聞 之? (Zhuangzi)
 Zi du wu_j hu_i [v_P wen zhi [p_P t_i t_j]]?
 you alone whom from hear 3.obj
 ‘From whom did you alone hear it?’



With the Intervention Effect of negation, the complex *wh*-complement 何所 *he suo* ‘what place’ in (16a) (=14) moves from its postverbal base position to the specifier of

HighFocP. Since there is a null preposition, the head of HighFocP is not occupied by ZHI. The tree diagram of (16a) is (16b).

- (16)a. 子 何 所 不 逞 欲? (Zuozhuan)
 Zi [he suo]_i bu cheng yu [pp ti]?
 you what place not satisfy desire
 '(In) what place do you not satisfy desires?'



4. Conclusion

In this paper I explore the Intervention Effect of negation on *wh*-adverbials in LAC. Since LAC is a *wh*-fronting language, *wh*-complements of adverbials must front to a higher position in the medial domain. There are two focalised positions between TP and *v*P for preposed *wh*-constituents: the High focus position above negation is expected to exclusively accommodate *wh*-complements of ‘high’ reason adverbials base-generated above negation, yet the Low focus position below negation allows *wh*-adverbials base-generated between negation and *v*P or postverbally. When a negator is present and c-commands a *wh*-adverbial that is base-generated between negation and *v*P or postverbally, *wh* cannot land in the Low focus position c-commanded by negation, as its Q-binding would be blocked by the negative barrier. Therefore, negation triggers *wh*-movement to the High focus position that is expected to allow ‘high’ PPs exclusively.

Of course, there are still remaining issues for future research such as: the limited possibilities of *wh*-in-situ, the motivation for *wh*-fronting, the precise nature of the Intervention Effect for focus constructions, and the presence/absence of the Intervention Effect of negation in LAC compared to modern Mandarin.

References

- Aldridge, E. (2010). Clause-internal *Wh*-movement in Archaic Chinese. *Journal of East Asian Linguistics* 19, 1-36.
- Aldridge, E. (2013). Object Relative Clauses in Archaic Chinese. *Canadian Journal of Linguistics* 58.2, 239-265.
- Beck, S. (1996). Quantified structures as barriers for LF movement. *Natural Language Semantics* 4, 1-56.
- Beck, S. & Kim, S-S. (1997). On *wh*- and Operator Scope in Korean. *Journal of East Asian Linguistics* 6, 339-384.
- Djamouri, R. (2001). Markers of predication in Shang bone inscriptions. In Chappell, H. (ed.), *Synchronic and Diachronic Perspectives of the Grammar of Sinitic Languages*, 143-171. Oxford: Oxford University Press.
- Djamouri, R. & Paul, W. (2009). Verb-to-preposition reanalysis in Chinese. In Crisma, P. & Longobardi, G. (eds.), *Historical syntax and linguistic theory*, 194-211. Oxford: Oxford University Press.
- Djamouri, R. Paul, W. & Whitman, J. (2012). Syntactic change in Chinese and the argument-adjunct asymmetry. In Cao, G., Chappell, H., Djamouri, R. & Wiebusch, T. (eds.), *Breaking down the barriers: Interdisciplinary studies in Chinese linguistics and beyond*, 577-594. Taipei: Academia Sinica.
- Feng, S. (1996). Prosodically constrained syntactic changes in Early Archaic Chinese. *Journal of East Asian Linguistics* 5, 323-371.
- Hagstrom, P. (1998). Decomposing Questions. Ph.D. dissertation, Massachusetts Institute of Technology.
- Kim, S-S. (2002). Intervention Effects Are Focus Effects. In Akatsuka, N. & Strauss, S. (eds.), *Japanese/Korean Linguistics* 10, 615-628. Stanford: Center for the Study of Language and Information.
- Kim, S-S. (2006). Intervention Effects in Questions. Ph.D. dissertation, University of Frankfurt.
- Meisterernst, B. (2010). Object Preposing in Classical and pre-Medieval Chinese. *Journal of East Asian Linguistics* 19, 75-102.
- Paul, W. (2002). Sentence-internal topics in Mandarin Chinese: the case of object preposing. *Language and Linguistics* 3.4, 695-714. Taipei: Academia Sinica.
- Paul, W. (2005). Low IP area and left periphery in Mandarin Chinese. *Recherches linguistiques de Vincennes* 33, 111-134.
- Pesetsky, D. (1999). *Phrasal Movement and Its Kin*. Cambridge: MIT Press.
- Peyraube, A. (1996). Recent Issues in Chinese Historical Syntax. In Huang, C.-T. J. & Li, Y-H. A. (eds.), *New Horizons in Chinese Linguistics*, 161-214. Dordrecht: Kluwer.

- Peyraube, A. (1997). On word order in Archaic Chinese. *Cahiers de linguistique Asie Orientale*, 26 (1), 3-20.
- Peyraube, A. (2003). On Moving Constituents in Chinese Historical Syntax. *Historical Development of Chinese Language*, 141-156.
- Wang, A. (2013). Object Preposing in Late Archaic Chinese. *York Papers in Linguistics* 13, 70-98.
- Wang, L. (1958). *Hanyu shigao*. Reprinted in 2004. Beijing: Zhonghua Shuju.
- Xu, D. (2006). *Typological Change in Chinese Syntax*. Oxford: Oxford University Press.
- Zhou, F. (1963). Stages in the development of the Chinese Language. In Zhou, F. (ed.), *Zhongguo yuwen luncong*, 432-438. Taipei: Zhengzhong Shuju.

Aiqing Wang
Department of Language and Linguistic Science
University of York
York
YO10 5DD
UK

aiqing.wang@york.ac.uk