PROCEEDINGS OF GLOW IN ASIA IX 2012:
THE MAIN SESSION

edited by
Nobu Goto, Koichi Otaki, Atsushi Sato, and Kensuke Takita
Proceedings of GLOW in Asia IX 2012:

THE MAIN SESSION

Edited by Nobu Goto, Koichi Otaki, Atsushi Sato & Kensuke Takita

Publication Date: March 27, 2013
This volume contains oral papers presented at the GLOW-in-Asia IX held at Mie University, Japan on September 4-6, 2012. Following the success of the first GLOW-in-Asia Workshop for Young Scholars at Mie University in the previous year, we were again extremely fortunate to have received a substantial number of abstracts from all over the world. A total of 18 oral and 32 poster presentations were selected and presented during the three-day conference, all of which were of high quality and represented most up-to-date research findings in a wide range of subfields in generative linguistics. In addition to the refereed presentations, we were privileged to have seven distinguished scholars to present their research: Anders Holmberg (Newcastle University), Junko Ito (UC Santa Cruz), Armin Mester (UC Santa Cruz), Uli Sauerland (ZAS) and Akira Watanabe (University of Tokyo) for the main session, and Hiromu Sakai (Hiroshima University) and William Snyder (University of Connecticut) for the workshop on experimental linguistics. The conference was well attended by over 100 participants and we had excellent discussions throughout the three days. We would like to acknowledge generous financial support from Mie University; a special mention is due for President Atsumasa Uchida and Vice President Hye-Sook Park for their understanding and support for the workshop in 2011 and the conference in 2012. We also wish to thank Prof. Mamoru Saito at Nanzan University and Prof. C.-T. James Huang at Harvard University and other executive members of the GLOW-in-Asia for having shared their expertise and assistance in the organization of the GLOW-in-Asia IX. Finally, our deepest gratitude is expressed to the abstract reviewers who contributed to the overall quality of the workshop.

The GLOW-in-Asia at Mie Local Organizing Committee

Chair: Seiki Ayano
Vice-chairs: Miki Obata, Koji Sugisaki
Members: Nobu Goto, Noriko Hattori, Koichi Otaki, Osamu Sawada, Kohei Suzuki, Kensuke Takita, Tomoko Tani, Etsuko Yoshida
# CONTENTS

1. **Anders Holmberg**  
   *The Syntax of Negative Questions and Their Answers* ....................................................... 1-18

2. **Mary Goodrich & William Snyder**  
   *Atelic Paths and the Compounding Parameter: Evidence from Acquisition* ............ 19-30

3. **Uli Sauerland & Jonathan David Bobaljik**  
   *Syncretism Distribution Modeling: Accidental Homophony as a Random Event* ....................................................... 31-53

4. **Akira Watanabe**  
   *Uninterpretable Features and Agreement* ........................................................................ 55-75

5. **Samuel Epstein, Hisatsugu Kitahara & T. Daniel Seely**  
   *Simplest Merge Generates Set Intersection: Implications for Complementizer ‘Trace’ Explanation* ................................................................................... 77-92

6. **Richard Faure**  
   *Resolutive Predicates and the Syntactic and Semantic Selection of Questions* ....................................................... 93-108

7. **Haoze Li**  
   *Association between Focus Particles and Interrogative Wh-Phrases* .................... 109-124

8. **Na Liu**  
   *The Syntactic Structures of the Chinese Bei Passive and the English Be-passive* ................................................................................... 125-139

9. **David J. Medeiros**  
   *A Cyclic Linearization Approach to Polynesian VP-remnant Movement* ................ 141-155

10. **Mythili Menon**  
    *The Apparent Lack of Adjectival Category in Malayalam and Other Related Languages* ................................................................................... 157-171

11. **Troy G. Messick**  
    *Ellipsis and Reconstruction in Tough Infinitives* ......................................................... 173-185

12. **Dan Michel & Grant Goodall**  
    *Finiteness and the Nature of Island Constraints* ......................................................... 187-197

13. **Chizuru Nakao, Masaya Yoshida & Ivan Ortega-Santos**  
    *On the Structure of Japanese ‘Why’-Stripping* ......................................................... 199-212
14. Chieu Nguyen  
Fusion and Scattering in Particle Collisions: When Universal Quantification and Contrastive Focus Interact with Themselves and Each Other in Vietnamese —— 213-228

15. Yoshiki Ogawa & Fumikazu Niinuma  
On the Syntactic Licensing of Locative Expressions in Japanese ———— 229-244

16. Shin-Ichiro Sano  
Patterns in the Avoidance of Marked Segmental Configurations in Japanese Loanword Phonology ———————————————————— 245-260

17. Ana Paula Scher  
Concatenative Affixation in Brazilian Portuguese Truncated Forms ———— 261-270

18. Yanan Sheng & Fuyun Wu  
Audience Design Affects Classifier Positioning in Chinese Relative Clause ——— 271-282

19. Yu Tanaka & Jun Yashima  
Deliberate Markedness in Japanese Hypocoristics ———————————————————— 283-297
1. Introduction

There is a well-known distinction among the languages of the world regarding how negative polar questions (yes/no-questions) are answered. This can be illustrated by the following examples, from Japanese and English, respectively (Q = question, A = answer):

(1) Q: Kimi tukarete nai? [Japanese]
   you tired NEG
   ‘Are you not tired?’
A: Un, tukarete nai.
   yes tired NEG
   (Lit.) ‘Yes, I’m not tired.’

(2) Q: Are you not tired? [English]
A: No, I’m not tired.

When confirming the negative alternative (I’m not tired), Japanese speakers use the particle which is used for affirmative answers to neutral questions, while English speakers use the particle used for negative answers to neutral questions. Correspondingly, when confirming the positive alternative of a negative question (denying the negative alternative), Japanese speakers use the particle otherwise used for negative answers to neutral questions, while English speakers use the particle otherwise used for affirmative answers to neutral questions.

(3) Q: Kare-wa kooihii-o nomai no? [Japanese]
    he-TOP coffee-ACC drink NEG Q
    ‘Does he not drink coffee?’
A: Uun, nomu yo.
    no drink PRT
    (Lit.) ‘No, he drinks (coffee).’

(4) Q: Does he not drink coffee?
A: Yes, he does.

This is the simple picture, to be modified in the course of this paper (see Kuno 1978, Pope 1976, Jones 1999: 4-14), Holmberg, in press). The two answering systems have been referred to in the literature as the truth-based and the polarity-based answering systems (Jones 1999). The truth-based system is what we see in Japanese. The logic of the nomenclature is that the

* I wish to thank the organisers of GLOW in Asia IX, Seiki Ayano and Koji Sugisaki, for inviting me, and the audience at the conference for their interesting comments and questions. Thanks also to Laura Bailey, Alex Leung, Somphob Yaisomanang and Ayaka Sugawara, and to Hilda Koopman and Craig Sailor for help with setting up the investigation on SSWL.
negative question is, as it were, about the truth of the negative proposition (‘Is it true that you are not tired?’), and the answer particle confirms or disconfirms it: ‘Yes, it’s true that I’m not tired’, or ‘No, it’s not true that I’m not tired.’ (i.e. ‘I am tired’). In the polarity-based system, on the other hand, the answer particle reflects, or agrees with, the polarity of the proposition in the answer. If the proposition is negative (confirming the alternative that I’m not tired), the answer particle is the negative no, if the proposition is positive, the particle is the positive yes.

Another nomenclature is that the Japanese system is an agree/disagree system: The answer conveys whether the respondent agrees or disagrees with the negative expectation conveyed by the negative question (cf. Pope (176: 73), Kuno 1978). Un in (1) would mean ‘I agree with your expectation that I’m not tired’. Neither nomenclature is ideal. However, for reasons to be made clear later I prefer the truth-based/polarity-based alternative.

The distinction seems to be generally regarded as a matter of cultural conventions, comparable to, say, the distinction between shaking hands or making a bow when greeting a new acquaintance. The alternative is that it is a matter of syntax: something about the syntactic rules of a language dictates whether it will employ the truth-based or the polarity-based answering system. In other words, it would be a case of parametric variation. There are some indications that the cultural convention hypothesis may be right. One of them is that there appears to be some intra-linguistic, possibly idiolectal variation in the use of ‘yes’ and ‘no’. Nevertheless, I will argue in this paper that the choice of ‘yes’ or ‘no’ when answering a negative question is fundamentally a syntactic matter, determined by the syntax of negation in the question, which is subject to parametric syntactic variation. English turns out to be a particularly interesting language in this connection. It has several syntactically distinct negations, and exhibits systematic variation with regard to answers to negative questions which correlates with the choice of negation in the question. The theory constructed on the basis of English will be shown to make predictions for other languages which do receive some cross-linguistic confirmation.

2. The Cross-linguistic Distribution of the Two Systems

The following is a list of languages reported to employ the truth-based system and languages reported to employ the polarity-based system. The list is based on descriptive grammars, on data gathered with the help of Syntactic Structures of the World’s Languages (SSWL), and from fieldwork. Each class includes only languages from different genera. The genus is named within the brackets as well as the family/phylum in cases where there are several languages from the same family/phylum. The location is mentioned where relevant.

**Truth-based**
- Afrikaans (Germanic, South Africa)
- Amele (Gum, Trans-New Guinea)
- Amharic (South Semitic, Afro-Asiatic)
- Mandarin (Chinese, Sino-Tibetan)
- Evenki (Tungusic)
- Georgian (Kartvelian)
- Ibibio (Lower Cross, Niger Congo)
- Japanese (Japonic)
- Kashmiri (Dardic, Indo-European)
- Kobon (Kalam-Kobon, Trans-New Guinea)
- Korean (isolate, East Asia)
- Kuot (isolate, Papua New Guinea)
- Lao (Lao-Putai, Tai-Kadai)
- Malagasy (Barito, Austronesian)

- 2 -
The two classes are not strictly parallel. The truth-based class includes some languages reported to allow both types of answers, while the polarity-based class includes only languages which are reported not to allow truth-based answers.\(^1\) Some areal effects are prominent in the list: There are no European languages in the truth-based class. East Asian languages are all in the truth-based class. East Asian languages are all in the truth-based class. Papua-New-Guinean languages are also all in the truth-based class. Arabic (varieties of) (Semitic, Afro-Asiatic)

---

\(^1\) Of the 33 languages in the SSWL with data about answering systems (at the time of writing), 16 are reported as allowing polarity-based answers only, 6 are reported as allowing truth-based answers only. 7 languages are reported as allowing both types. They are Afrikaans, Ibibio, Japanese, Mandarin, Nweh, Taiwanese Southern Min, and Georgian. Kashmiri is reported to allow both types in Wali and Kouli (1997), as is Mauwake, in Berghäll (2006). For some languages I have contradictory results. Hausa (Chadic, Afro-Asiatic) is reported to follow the truth-based system in SSWL, but to follow the polarity-based system in my own fieldwork with an informant.
truth-based class, while African languages are found in both classes. This does not have any clear implications for the choice between the ‘cultural hypothesis’ and the ‘syntactic hypothesis’. Cultural traits obviously spread through contact between peoples, but so do syntactic features. One of the most important findings of linguistic typological research from the last twenty years is that linguistic features can be geographically distributed over surprisingly large, even continent-wide areas, across family lines (Dryer 1998). I will not say more about the typological aspect of answering systems in this paper.

3. A Language which Allows both Options: English

One language which appears to allow both the truth-based and the polarity-based option is English. Consider (4) and (5). The question in (5) should be read as conveying an expectation that John is coming. For example, imagine that Bill says to Liz and Mary “Do you know who is coming along on this trip”. Mary then turns to Liz and asks (5). The answer Yes then means ‘John is coming’, while No means he is not coming; the polarity-based system.²

(5) Q: Isn’t John coming?
A: Yes. (‘John is coming.’)
A: No. (‘John is not coming.’)

The question in (6) would normally convey the expectation that John is not coming. Imagine, for example, that Bill eyes the group of people who have turned up for the trip, and notices that John is not among them, at which point he asks the question in (6).

(6) Q: Is John not coming?
A: Yes. (‘John is not coming.’)
A: No. (‘John is not coming.’)

For many speakers of English the answer yes now conveys confirmation that John is not coming. This looks like the truth-based system. The answer no also conveys confirmation that John is not coming, though, which is not characteristic of the truth-based system. This means that yes and no mean the same thing in the context of (6). Kramer and Rawlins (2011, 2012) refer to this as negative neutralization. So in this case English is half-way truth-based, as it were. Below it will be shown that there are contexts where English replicates the truth-based system exactly.

(5) and (6) imply that choice of negation in the question determines what the answer yes means. Judging by just (5) and (6), choice of n’t means that yes confirms the positive alternative, while choice of not means that yes confirms the negative alternative. In the following I will argue for the following hypothesis, call it Hypothesis 1.

(7) Hypothesis 1: The choice of answering system in a language, depends on the syntax of polarity, including the syntax of negation.

It will be shown that Hypothesis 1 can best be understood if Hypothesis 2 holds true (see Kramer and Rawlins (2011, 2012), Holmberg (in press)):

² For a subset of native English speakers this is the only reading (2) can have, while for the other speakers it can, in the right context, convey expectation of a negative answer. I will return to this distinction below in section 4.
(8) **Hypothesis 2**: Answers like ‘yes’ and ‘no’ are complete sentential expressions, where the IP is elided (not spelled out), under identity with the IP of the question.

Even though many native speakers of English agree that *yes* in the context of (6) means that John is coming, there are also many who do not agree with this interpretation. For these speakers the bare answer ‘yes’ in (6) is not a well formed answer. The intuition of these speakers is that it is indeterminate: *Is he coming or isn’t he?*

(9) Q: Is John not coming?  
A: #Yes. [some speakers of English]  
A: Yes he is.

An unquestionably well-formed alternative is the longer version *Yes he is*, accepted by all English speakers as unambiguously meaning (in this case) that he is coming, disconfirming the negative alternative.

Yet another complication is that we also find (10), where the answer disconfirms the negative alternative, just like *Yes he is* in (9).

(10) Q: Is John not coming?  
A: No, he is.

We can summarise the various alternatives for answering negative questions in English as follows:

(11) Q: Is John not coming?  
A: Yes.  
    Some speakers: ‘John is not coming.’  
    Other speakers: Not a well formed answer.  
A: Yes he is.  
    All speakers: ‘John is coming.’  
A: No.  
    All speakers: ‘John is not coming.’  
A: No, he is.  
    Some (or all) speakers: ‘John is coming.’

The upshot is that English does exhibit the conjunction of the possibilities allowed by the truth-based system and the polarity-based system, if we include all the varieties of English discussed above. We may note that we do not, at present, know how the two or three different varieties are distinguished in social or regional terms.

There is another point where there is variation among speakers of English. For many speakers, (5) (*Isn’t John coming?*) unambiguously conveys expectation of a positive answer, while for other speakers it is ambiguous out of the blue: Depending on the context, it may convey a positive or a negative expectation. Ladd (1981), discussing the ambiguity of

---

3. Kramer and Rawlins (2012) reported a carefully made experiment demonstrating this variation among speakers of English in relation to negative questions with *not.*

4. *No, he is* in (9) is spelled with a comma, while *Yes he is* in (8) is spelled without one. This reflects the observation that *Yes he is* can be, and perhaps typically is, pronounced as a single intonation unit, while *No, he is* in (8) is typically pronounced as two intonation units.
negative questions such as (5) in the latter variety of English observes that the reading can be controlled for by polarity items.\(^5\)

(12) a. Isn’t John coming, too?
   b. %Isn’t John coming, either?

(12a) can only convey a positive expectation, while (12b), for the speakers who accept it at all (the ‘ambiguous variety’ for whom (5) is ambiguous), it conveys a negative expectation. For the other speakers (the unambiguous variety) it is plain ungrammatical. Ladd (1981) argues that this is explained by the scope of the negation: For the ambiguous variety, \(n’t\) moved to C may have low scope, licensing the negative polarity item either. This point will become clearer below after discussing the structure of questions and answers in more detail.

Now, in the ambiguous variety, plain \textit{yes} is not a well-formed answer in a question with \(n’t\) where \(n’t\) has low scope, while \textit{yes}+V and (at least for many speakers \textit{no}+V) is a well formed answer.

(13) Q: Isn’t John coming, either?
   A: #Yes.
   A: Yes he is.
   A: (%)/No, he is.

4. The Three Negations of English

Consider the following question-answer triplet (discussed in Holmberg, in press).

(14) Q: Does John \textit{sometimes} not show up on time for work?
   A: Yes. (‘John sometimes does not show up on time for work.’)
   A: No. (‘John does not sometimes not show up on time for work’, i.e. ‘He is always on time’.)

In this case, where an adverb is inserted before the negation in the question, the answer \textit{yes} unambiguously, and for all speakers of English, confirms the negation, i.e. means that John sometimes does not show up for work.

The answer \textit{no} is a bit harder to process, but once processed, the meaning is the one given within brackets, denial of the negative alternative, meaning that John is always on time. The reason why it is harder to process is, quite clearly, the (understood) double negation.

So in this case English conforms strictly to the truth-based system. (15) and (16) are two additional examples illustrating the same point: insertion of an adverb before the negation forces a truth-based reading of \textit{yes} and \textit{no}.

(15) Q: Did you \textit{purposely} not dress up for this occasion?
   A: Yes. (‘I purposely did not dress up.’)
   A: No. (‘I did not purposely not dress up; I just didn’t know the dress code.’)

(16) Q: Do cats \textit{typically} not like rotten food?
   A: Yes. (‘They typically don’t like rotten food.’)

\(^5\) Ladd is apparently not aware that he is describing a restricted variety of English. I have the impression that the ‘unambiguous variety’ is more common in the USA than in Great Britain. But Robert Ladd himself is American, and I have encountered speakers of British English who consider (5) to be unambiguous.
A: No. (‘They don’t typically not like rotten food; typically they don’t mind if the meat is a bit rotten.’)

In (17) I have summarized the findings up to now with regard to the meaning of yes and no as answers to negative questions in English. The focus particle too is added in the question (17a) to force the positive expectation reading.

(17a) Q: Isn’t John coming, too?
Aa: Yes.
No speaker variation: ‘John is coming.’

b. Q: Is John not coming?
A: Yes.
Speaker variation: ‘John is not coming’ or an indeterminate, infelicitous answer.

c. Q: Is John usually not coming?
A: Yes.
No speaker variation: ‘John is usually not coming.’

It is well known that there is a syntactic difference between the negations n’t and not in English, most clearly seen under subject-auxiliary inversion (T-to-C), where n’t follows the moved auxiliary but not does not. But there are also two negations not. This is, basically, the reason behind the variation in the meaning of the answers yes and no to negative questions. The existence of two negations is uncontroversial (see Cormack and Smith 2002): They can co-occur in the same sentence.

(18) You cannot not go to Church, and still call yourself a good Christian.
(19) You must not ever not address him as ‘Sir’.

I assume the structure is as in (20).

(20)

I will refer to the higher not as middle not (implying that there is a still higher one, to be discussed below) and the other not as low not. Middle not has sentential scope; I will return below to how this scope is established. Low not has scope over VP only. An obvious difference between them is that middle not can be substituted by n’t, low not cannot.

(21) a. You mustn’t ever not address him as ‘Sir’.
   b. *You must not evern’t address him as ‘Sir’.
Now consider again the question *Is John not coming?*. Following Holmberg (in press), I put forward the linked hypotheses in (22):

(22) a. When the question is analyzed (parsed) as having low negation, the answer *Yes* means ‘John is not coming’, and  
b. when the question is analyzed (parsed) as having middle negation, the answer *Yes* is not a well formed expression.

To understand how this works we need to first discuss the syntax of questions.

6. The Syntactic Structure of Questions and Answers

I assume that all questions have the basic structure (23):

(23) \[ Q \ x \ Foc \ \left[ IP \ldots x \ldots \right] \]

There is a free variable which is the focus of the question, the focusing derived by movement of the variable to specFocP. In direct questions there is, in addition, an illocutionary force feature \( Q \) which encodes a request to the addressee to provide a value for the variable such that the resulting proposition is true. In wh-questions the variable is a wh-phrase, overtly moved to spec,FocP in English and other languages with wh-movement. The answer provides a value for the variable. In yes/no-questions the variable is polarity. The basic structure of *Is John coming?* is (24).

(24)

\[
\begin{align*}
&\text{Q} \\
&\text{uPol} \\
&\text{Foc} \\
&\text{IP} \\
&\text{John [uPol] is coming}
\end{align*}
\]

The variable is a formally unvalued feature [uPol] which has two possible values, [+Pol] and [-Pol]. The answer, either *Yes* or *No*, provides a value for the variable.\(^6\) The structure of the answer is (25).

(25)

\[
\begin{align*}
&\text{yes} \\
&\text{[+Pol]} \\
&\text{Foc} \\
&\text{IP} \\
&\text{John [+Pol] is coming}
\end{align*}
\]

The IP of the answer is identical to that of the question except that the polarity variable is assigned a value by the particle merged in spec of Foc. This is just the notion of identity

---

\(^6\) See Bailey (2012) on the difference between questions with and questions without illocutionary question force, including (most) embedded questions. The hypothesis argued for by Bailey is that the formal difference is the presence of the Q-feature.
required for ellipsis. It is well known that strict referential identity is not required for ellipsis; compare a case like (26), exemplifying VP-ellipsis.

(26) John, doesn’t hate himself, but Bill does (hate himself).

The VP in the second conjunct is identical to the VP in the first conjunct except that the referential value of the reflexive, a referentially variable item, is different. Following Holmberg (in press) I postulate (27).

(27) A syntactic constituent \( \alpha \) can be elided if it has a discourse-local antecedent which is identical with \( \alpha \) at LF, up to assignment of values to variables.\(^7\)

Typically the antecedent is found in a higher clause, or in the preceding independent sentence in a discourse. I leave the precise meaning of ‘discourse-local’ open.

Being identical up to the assignment of values to variables with an antecedent in the immediately preceding question, the IP of the answer (25) can be, and usually is elided, i.e. not pronounced. Alternatively the merged, focused particle is \( \text{no} \), encoding \([-\text{Pol}]\), assigning negative value to \([\text{uPol}]\) in IP. Again, the IP is typically not spelled out. If it is spelled out it will be \( \text{No, John is not coming or No, John isn’t coming,} \) with \([-\text{Pol}]\) spelled out as \( \text{not or n’t.} \)

Consider the structure of a question with low negation (28a), where the adverb is included to force the low reading of the negation; by hypothesis the negation can be merged low regardless whether there is an adverb, though. The affirmative answer in this case gets the truth-based interpretation, (28b).

(28) a. Is John sometimes not coming (to work)?
   b. Yes. (‘John is sometimes not coming.’)

The structure of the question is (29a), and the structure of the affirmative answer is (29b), where most often all that is pronounced is the focused affirmative particle.

(29) a. \([\text{is, uPol}] \text{ Foc [IP John [is, uPol] [sometimes [VP not coming ]]]}\]
   b. \([\text{yes, +Pol}] \text{ Foc [IP John [is, +Pol] [sometimes [VP not coming ]]]}\]

The answer affirms the negative alternative (that John is sometimes not coming). This is the truth-based type of answer.

The negative answer will have the structure (30), the polarity variable assigned negative value by the focused negative particle.

(30) \([\text{no –Pol}] \text{ Foc [IP John [is, –Pol] [sometimes [VP not coming ]]]}\]

Since the negation is inherently valued \([-\text{Pol}], \) the sentence will have double negation. The IP is normally elided, leaving just the focused negative particle spelled out. The interpretation is ‘No, John is not sometimes not coming’, i.e. ‘No, he is always coming (to work)’, the truth-based reading.

This explains (22a). Now consider the answer of a negative question with middle negation. As stated in (22b), in this case (bare) \( \text{yes} \) is not a well formed answer. The general idea is that this is because middle negation in the IP of the answer clashes with the positive

---

\(^7\) See Merchant (2001). See Krifka (2012) and Holmberg (in press) for an argument that just semantic identity (mutual entailment) is not a sufficient condition for ellipsis in question-answer pairs (hence the qualification “identical with \( \alpha \) at LF”).
polarity value of the focused polarity particle *yes*; they are, as it were, too close to one another. To work this out formally, we need to consider the more precise role of the polarity head in IP.

I assume that a declarative sentence, for instance *John is coming*, has the (simplified) structure (31): There is a polarity head, by assumption the highest head in the IP-domain, which is always merged unvalued. If the sentence does not contain a negation or any other constituent with inherent negative polarity value, the polarity head is valued positive by default. I continue using the label ‘IP’, although it is by assumption headed by Pol, hence a PolP. The auxiliary is adjoined to [uPol] through head-movement.

(31)  \[ C \mid [IP \mid [is, \textbf{uPol}] \mid [VP \textit{coming}]] \rightarrow C \mid [IP \mid [is, \textbf{+Pol}] \mid [VP \textit{coming}]] \]

If, however, the sentence contains a negation close enough to the unvalued polarity head, it will assign negative value to it. This is the case of middle negation, as in *John is not actually coming*, where the adverb is included to ensure a middle negation reading. The derivation is as depicted in (32).

(32)  \[ C \mid [IP \mid [is, \textbf{uPol}] \mid [\textbf{not} \mid [\textit{actually} \mid [VP \textit{coming}]]]] \rightarrow C \mid [IP \mid [is, \textbf{–Pol}] \mid [\textbf{not} \ldots]] \]

[Pol] and middle negation form a chain sharing an interpretable [–Pol] feature (so this is not a case of double negation).

In a yes/no-question, the sentential polarity head is unvalued, as discussed earlier, and focused, by movement to the spec of Foc in the C-domain. This is, by hypothesis, the case whether or not there is a negation in the sentence. We thus assume that focusing of polarity precludes internal valuation of [uPol] by an inherently valued negative item. The structure of a negative question (33a) with middle *not* is (33b):

(33)  a.  Is John not actually coming?
b.  \[ [is, \textbf{uPol}] \text{Foc} \mid [IP \mid [is, \textbf{uPol}] \mid [\textbf{not} \mid [\textit{actually} \mid [\textit{coming}]]]]] \]

In the answer, a declarative sentence with a focused polarity particle, there are two items with inherent polarity value competing to assign a value to [uPol] in IP, the middle negation and the affirmative polarity particle.

(34)  \[ [CP \mid [yes, \textbf{+Pol}] \text{Foc} \mid [IP \mid [is, \textbf{uPol}] \mid [\textbf{not} \mid [\textit{actually} \mid [\textit{coming}]]]]] \]

The result is the indeterminate reading seen in (11) and (17b).

The speaker variation in (17b) (also illustrated in (11) and (13)) can now be understood as follows: Some speakers of English assign a low negation reading to *not* in a polar question as their default reading (or preferred reading). A low, VP-internal negation is distant enough from the sentential polarity head not to assign value to it, meaning that the affirmative particle can assign positive value to [uPol] uncontested.

(35)  \[ [CP \mid [yes, \textbf{+Pol}] \text{Foc} \mid [IP \mid [is, \textbf{+Pol}] \mid [VP \textbf{not} \mid [\textit{coming}]]]]] \]

These speakers readily answer *yes* to a negative sentence with *not*, even out of context, to confirm the negative alternative. Other speakers assign a middle negation reading to *not* as their default (or preferred) reading in polar questions, hence cannot answer with *yes* to confirm the negative alternative of a negative question.
As was illustrated in (13), a subset of native speakers of English can use (and understand) a question with \( n't \), such as (36), to convey expectation of a negative answer. The reading is forced in (36) by the NPI; for other speakers (36) with the forced negative expectation reading is ungrammatical.

(36) Isn’t John coming (either)?

This can be understood as follows: In the variety of English which accepts (36), \( n't \) is moved from IP, more precisely, from the middle negation position in IP, to the C-domain. The negative chain can be interpreted as having the interpretable link inside IP, in the middle negation position. In that case, the sentence licenses an NPI in IP and conveys a negative expectation. Alternatively, only the higher link is interpretable negative, in which case the sentence conveys a positive expectation, and excludes an NPI in IP. For a subset of English speakers this is the only alternative. Effectively, in this variety, \( n't \) is externally merged in the C-domain in polar questions. I will refer to the high reading of \( n't \) as high negation (contrasting with middle and low negation, already discussed).

As shown in (11), repeated here as (37), all speakers of English can answer a negative question with \textit{yes} followed by a subject and an auxiliary but with elided VP.

(37) Q: Is John not coming?  
A: Yes he is.  
All speakers: ‘John is coming.’

This is because in this case only the VP is elided, hence only the VP needs to have an identical antecedent in the question, so the fact that the question contains a middle negation does not constrain the answer, which is just a plain affirmative sentence confirming the positive alternative, contradicting the negative alternative posed by the question.\(^8\)

(38) \[
\begin{array}{c}
\text{[CP [yes, +Pol] Foc [IP he [is, +Pol] [VP coming ]]]}
\end{array}
\]

The negative neutralization which Kramer and Rawlins (2011) observed and discussed, shown in (6), repeated here, can now be understood as follows:

(6) Q: Is John not coming?  
A: Yes. (‘John is not coming.’)  
A: No. (‘John is not coming.’)

The answer \textit{yes} means ‘John is not coming’ when the question is taken to have low \textit{not}. The answer \textit{no} means ‘John is not coming’ when the question is taken to have middle negation.

How does the polarity-based system work? The difference shows in negative answers to negative questions. Consider, again, (39):

(39) a. Is John not coming?  
b. No. (‘John is not coming’)  
c. \[
\begin{array}{c}
\text{[CP [no, –Pol] Foc [IP he [is, –Pol] not [VP coming ]]]}
\end{array}
\]

The question, by hypothesis, has middle \textit{not} (since with low \textit{not} the result is affirmation of the negative alternative, the truth-based answer). The focused negative particle in the answer

\(^8\) This was discussed in Holmberg (2001) in relation to Finnish.
assigns $[-]$ value to the polarity feature in IP. In this case, although there is another interpretable negation in the sentence the result is not double negation. We know the negation in the answer is interpretable, because it is identical to the negation in the question, or else it could not be elided. We also know that the answer particle *no* is interpretable negative because it assigns negative value in answers to neutral questions. What we seem to get is a sharing of an interpretable negative feature between Pol and the middle negation. This, then, seems to be a property characteristic of the polarity-based system, present in some languages but not others.

So is the distinction between the truth-based and the polarity-based answering system a matter of syntax or cultural conventions? The mere fact that English employs both systems, depending on the scope of the negation in the question, makes it apparent that we are dealing with a syntactic, structure-dependent phenomenon. I will pursue the hypothesis that this is, in fact, the explanation of the cross-linguistic variation between these two systems. To put it simply, languages with a consistently truth-based system only have a low negation, languages with a consistently polarity-based system only have middle or high negation, while languages with a mix are like English in having variation between a high and a low negation. I will now discuss two other language, Swedish, which is consistently polarity-based, and Thai, which is consistently truth-based.

7. **Swedish: a Language without Low Negation**

Swedish has a robustly polarity-based answering system.

(40)  

<table>
<thead>
<tr>
<th></th>
<th>Q:</th>
<th>A:</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>Har Johan kommit? [Swedish] has Johan come</td>
<td>Ja./Nej. yes / no</td>
</tr>
<tr>
<td>b.</td>
<td>Har Johan inte kommit? has Johan not come</td>
<td>*Ja./Nej. ('Johan has not come.')</td>
</tr>
</tbody>
</table>

There is no ambiguity or (as far as we know) any speaker variation regarding the affirmative answer in (40): It cannot affirm the negative alternative. This follows if the Swedish negation is exclusively a middle negation. As a middle negation it will clash with the feature of the affirmative particle, the situation depicted in (35), for English. That Swedish does not have low negation is further confirmed by the observation that double negation is not possible; compare (41a) and English (18).

(41)  

<table>
<thead>
<tr>
<th></th>
<th>Q:</th>
<th>A:</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>*Du kan inte inte gå i kyrkan, ... [Swedish] you can not not go to Church</td>
<td></td>
</tr>
<tr>
<td>b.</td>
<td>Du kan inte avstå från att gå i kyrkan, ... you can not refrain from going to Church</td>
<td></td>
</tr>
</tbody>
</table>

Swedish does not have a negation with VP-scope. To express negation with VP-scope, Swedish has to use a verb with lexical negative meaning, such as *avstå* 'refrain'.

In this light, it may seem surprising that inserting an adverb in the question has the same effect on the answer as in English; compare (42) and (14).

(42)  

<table>
<thead>
<tr>
<th></th>
<th>Q:</th>
<th>A:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Har Johan nångång inte kommit i tid? [Swedish] has Johan any time not come on time</td>
<td></td>
</tr>
</tbody>
</table>
A: Ja.  ('He has sometimes not been on time.')
A: Nej. ('He has not sometimes not been on time, i.e. he has always been on time.')

Following Holmberg (in press) this is explained as follows: The effect of the adverb preceding, i.e. c-commanding, the middle negation in the question, and thus, by hypothesis, also in the answer (even though it is usually not spelled out) is that the negation is prevented from valueing [uPol].

\[
(43) \quad \text{[CP [ja, +Pol] Foc [IP Johan [har, +Pol] [nångång [inte [kommit i tid]]]]]}
\]

The result is that the focused answer particle values [uPol]. If the particle is affirmative, the effect is confirmation of the negative alternative, i.e. that Johan is sometimes not coming on time. If the answer particle is negative, the result is double negation, i.e. that Johan is not sometimes not coming, i.e. he is always coming on time. The latter reading is harder to process, due to the double negation.

The effect of the lack of a low, VP-internal negation in Swedish, apart from the fact that it must resort to negative lexical verbs to express VP-negation, as in (41), is that the affirmative answer particle can never confirm the negative alternative of a negative question, the way it can in English, except when an adverb precedes the negation, thus intervening between the negation and the unvalued polarity head. In Kramer and Rawlins’s (2011) terms, Swedish does not have negative neutralization.

To confirm the positive alternative of a negative question, Swedish uses a special affirmative particle.

\[
(44) \quad \text{Q: Har Johan inte kommit? [Swedish]}
\]

\[
\text{A: Jo. yes.REV 'Yes he is coming.'}
\]

The particle is glossed as ‘yes.REV’, short for ‘affirmative polarity-reversing’ (see Farkas and Bruce 2009, Holmberg 2003). The effect of the particle is to neutralize the negative feature of the negation in the answer, or, in slightly different formal terms, reverse the negative polarity caused by the negation in the answer. Recall that English resorted to VP-ellipsis to resolve the same problem, namely, how to disconfirm (deny, negate) the negation inherited from a question with middle negation. As shown in Holmberg (2001), Finnish is like English in this respect. Like Swedish are the other Scandinavian languages, as well as German, Dutch, French, and Standard Arabic.

8. **Thai: a Language with Low Negation Only**

The standard polar question form in Thai employs the question particle mái (or mày).

\[
(45) \quad \text{Q: phiì-chaay pay paa-rïit mái [Thai]}
\]

\[
\text{older-brother go Paris Q}
\]

\[
\text{‘Did your brother go to Paris?’}
\]

\[
\text{A: pay go}
\]
The affirmative answer in this case is expressed by echoing the verb of the question (Yaisomanang 2012; see Jones (1999), Holmberg (2001, 2007) for the syntax of this form of answer). This question particle cannot be used with a negative question. The reason for this has to do with the fact that the question particle แปล is closely related to the negation แปล. According to Yaisomanang (2012) แปล is the spellout of รู้ แปล ‘or not’, and is therefore incompatible with a negation in the question (compare *Is John not coming, or not). To express a negative question Thai will use another question particle.

(46) Q: ผู้ชาย แปล ไป ปารีส [Thai]  
older-brother  NEG  go  Paris  Q  
‘Didn’t your brother go to Paris?’
A: ชาย ร ใช่ แปล  ‘No.’ (‘He didn’t go.’)

Here the affirmative particle ร ช ย unambiguously confirms the negative alternative of the question: the truth-based system.9

According to Yaisomanang (2012), the question particle รู้ is one of several spell-outs of ช่วย รู้ แปล lit. ‘right or not’.10 The structure of a negative question employing this particle is (47).

(47) Q-Force CP IP I I’ I 
 phi-ช่วย แปล ไป ปารีส [+Pol] VP ConjPol PolP 
brother  NEG  go  Paris  [−Pol] VP Conj’Pol PolP 
(chวย) รู้ (แปล ช่วย) ร ใช่ แปล รู้ แปล (ช่วย)
right  or  not  right

That is to say, the propositional content of the question is the subject of a predicate made up of the disjunction of a positive and a negative polarity phrase. The question particle is the spell out of a disjunctive predicate. This disjunctive predicate makes up the question variable:

9 Another way of confirming the negative alternative is by echoing the negation and the verb of the question.
A: แปล ไป
NEG  go
‘No.’ (‘He didn’t go.’)

This type of answer falls outside of the dichotomy between truth-based and polarity-based answering system.

10 Other spell-outs are ช่วย รู้, รู้ แปล, and ช่วย รู้ แปล (Yaisomanang 2012).
see (23) and (24). Q-Force tells the addressee to select the conjunct which, when predicated of the sentential subject yields a true proposition. The answer has the following structure:\(^{11}\)

\[
\begin{array}{c}
(+\text{Pol}) \ tch\text{-}y, \\
\text{Foc} \\
\hline \\
\text{IP} \\
\hline \\
\text{I} \\
\hline \\
\text{PolP} \\
\hline \\
\text{phii-chaay m\text{\-}y pay paa-riit} \\
\text{brother not go Paris} \\
\hline \\
[u\text{Pol}] \rightarrow (+\text{Pol})
\end{array}
\]

The focused positive polarity feature, lexicalized by \textit{tch\text{-}y}, assigns positive value to the polarity head in IP. The sentential subject is normally elided, under identity with the subject of the question. The negative answer would be (49), with \([-\text{Pol}, \text{tch\text{-}y}],\) spelled out \textit{m\text{\-}y tch\text{-}y}, in the focus position, assigning negative value to \([u\text{Pol}]\) in the (elided) matrix IP.

\[
\begin{array}{c}
m\text{\-}y tch\text{-}y. \\
\text{not right (`}He went to Paris.`')
\end{array}
\]

This is ‘low negation’ in the sense that there is no c-command between NEG and \([u\text{Pol}],\) the polarity variable, consequently there can be no feature valuation of \([u\text{Pol}]\) by the negation, and as a consequence of this, the answer is strictly truth-based.

9. Back to Japanese

Consider once more the Japanese example in (1), repeated here.

(1) Q: Kimi tukarete nai? \hfill [Japanese]  
you tired NEG  
`Are you not tired?'
A: \textit{Un}, tukarete nai.  
yes tired NEG  
(Lit.)`Yes, I’m not tired.’

If the theory up to now is right, the negation in the question, and therefore also in the answer, is low enough not to assign negative value to the \([u\text{Pol}]\) feature, assumed to universally head a finite sentence. The structure of the question would, very roughly, be (50a), and the structure of the answer, (50b), which should be compared with English (35): The negation is embedded within the predicate and as such does not (at least not necessarily) assign value to the sentential Pol head.

\[
\begin{array}{c}
\text{(50) a. } [\text{Q-Force [ [u\text{Pol}] Foc [IP DP}_1 [AP tukarete nai] [u\text{Pol}]]]]} \\
\quad \text{you tired NEG} \\
\text{b } [ \text{[un, +Pol] Foc [IP DP}_1 [AP tukarete nai] [+Pol]]] \\
\quad \text{yes I tired NEG}
\end{array}
\]

\(^{11}\) Here I am simplifying Yaisomanang’s analysis, for ease of exposition.
The truth-based answer in (1) is what we find when the question expects a negative answer. But a negative question in Japanese can also convey expectation of a positive answer, as in the following example.\textsuperscript{12}

\begin{tabular}{l}
(51) Q: Kore oisiku nai? [Japanese] \\
this delicious NEG \\
‘Isn’t this delicious?’
\end{tabular}

\begin{tabular}{l}
A: \textit{Un}, oisii. / \\
yes delicious
\end{tabular}

\begin{tabular}{l}
A: \textit{Uun}, oisii. \\
no delicious
\end{tabular}

Crucially, the intonation is different in this case. Very roughly, in a question with negative expectation, as in (1), there is a fall on the final negation, while in a question with a positive expectation, as in (51), the pitch remains high on the negation.

If the theory in this paper is right, this means that the negation in the question (51) is a high negation, outside the sentential [uPol] head. As such it does not reappear in the answer. Effectively, the negation has no effect on the answer. Compare the discussion of (36) above in English. The structure of the question is, very roughly, (52a), and the structure of the answer is (52b).

\begin{tabular}{l}
(52) a. [Q-Force [ [uPol] Foc [[[IP DP \textsubscript{1} \textsubscript{[AP tukarete] uPol ]] nai ]]]
\end{tabular}

\begin{tabular}{l}
this delicious NEG
\end{tabular}

\begin{tabular}{l}
b. [[un, +Pol] Foc [IP DP \textsubscript{1} \textsubscript{[AP tukarete] +Pol ]]]
\end{tabular}

The analysis of the negation as a high negation is consistent with the observation that this type of question cannot contain a question particle. (53) is a grammatical question, but unlike (51), it conveys expectation of a negative answer, and consequently the answer pattern is different.\textsuperscript{13}

\begin{tabular}{l}
(53) Q: Kore oisiku nai no? [Japanese] \\
this delicious NEG Q \\
‘Isn’t this delicious?’
\end{tabular}

\begin{tabular}{l}
A: \textit{Un}, oisii. / \\
yes delicious
\end{tabular}

\begin{tabular}{l}
A: \textit{Uun}, oisii. \\
no delicious (‘It is delicious.’)
\end{tabular}

The effect is that the form of the answer correlates with the expected answer: If the expected answer is negative, ‘yes’ confirms the negative alternative. If the expected answer is positive, ‘yes’ confirms the positive alternative. This is the ‘agree/disagree’ system, so called. If the theory here is right, however, the expectation conveyed by the question and the form/meaning of the answer both depend on the syntactic structure, specifically the scope of the negation in the question. There is no direct relation between expectation and form of answer. The ‘agree/disagree system’ nomenclature is, therefore, somewhat misleading.

\textsuperscript{12} I am indebted to Ayaka Sugawara for the examples and discussion of question and answer strategies in Japanese.

\textsuperscript{13} Ayaka Sugawara (p.c.) suggests ‘You don’t think this is delicious, do you?’ as the best English translation.
An interesting observation is that expected answer does not appear to affect the answer strategy in Thai. Consider the following example (provided by Somphob Yaisomanang, p.c.):

(53) Q:  thəə məy cha-ləat rûu  [Thai]
    she  NEG  clever
    ‘Isn’t she clever?’ (said by a mother about her daughter)
A:  chəy (thəə məy cha-ləat)
    yes  she  NEG  clever
    ‘No, she is not clever.’

A:  pləaw (thəə cha-ləat)
    NEG  she  clever
Lit:  No (she is clever).
i.e.  ‘Yes, she’s clever.’

This is a context where the expected answer is clearly confirmation of the positive alternative. Nevertheless, the answer follows the truth-based pattern. This follows if the analysis of the question in (53) is the same as the analysis of the question (46): the negation in the question remains out of reach of the sentential polarity head, regardless of what the speaker expects. In Japanese, on the other hand, as in English, the position of the negation is variable.

10. Conclusions and Consequences

I have argued in this paper that the distinction between the truth-based and the polarity-based answering systems visible in answers to negative polar questions is a matter of syntax, not cultural conventions (Hypothesis 1 in section 3). It depends on the position and scope of the negation in the question, which is crucial in answers, because the answer copies the IP of the question (although it is typically not spelled out; see Hypothesis 2 in section 3). The position of negation is subject to parametric variation. The languages which have the truth-based alternative have a negation which is low, in the sense that it is not accessible to the sentential Polarity head in answers to polar questions at all (as in Thai question and answer pairs), or is distant enough from the sentential unvalued polarity head to be trumped by the answer particle (as in English and Japanese, when the low negation is selected). The advantage, so to speak, of the truth-based system is that negative questions can be answered with the same answer particles that are used in answers to neutral questions. The polarity-based system has a negation which is high enough to interact with the Polarity head in answers to negative questions. In answers which are meant to confirm the positive alternative this leads to a feature conflict problem which is resolved in different ways; VP-ellipsis or a polarity-reversing affirmative particle are two solutions. In answers meant to confirm the negative alternative, feature-sharing between the negative particle and the negation is required to avoid double negation, in the polarity-based system.

The predictions that this makes for the two classes of languages listed in section 2 are clear: The languages which follow the truth-based system strictly have a negation which is always low (in the sense just described), while the languages which allow either system have more freedom as regards the scope of negation. The languages which follow the polarity-based system strictly do not have low negation, but a middle negation always accessible to the unvalued sentential polarity head (unless an adverb can intervene), or a high, IP-external negation. Testing these predictions will require considerable work. Simple inspection of word order in declarative or interrogative negative sentences can give clues, but is probably not sufficient, in most cases, to determine the scope of negation.
It was mentioned in the introduction that there are some indications that something more, or something other than syntax is involved. Informants speaking the same language sometimes disagree over how to answer negative questions, as I have found in my fieldwork. There are also anecdotal reports by parents that young children sometimes appear to follow the wrong system. For example, my own son followed the truth-based system when speaking his mother tongue Swedish at least until he was seven. Do these children really misanalyse the scope of negation in their language? If the theory articulated in this paper is right, that is what they do.

References

ATELIC PATHS AND THE COMPOUNDING PARAMETER: EVIDENCE FROM ACQUISITION*

Mary Goodrich and William Snyder
University of Connecticut

1. Overview

In an acquisitional study of motion predicates, Snyder, Felber, Kang, & Lillo-Martin (2001) reported a strong association between first novel compound words, and first atelic path phrases with motion verbs, in children acquiring English. Some of the children acquired novel compounding early, some late, but whenever a given child acquired compounding, atelic path phrases followed soon after.

This finding, if correct, has important implications for the nature of cross-linguistic variation in the expression of motion events. Yet, Gehrke (2008) has questioned Snyder et al.’s evidence, noting that many of their "atelic" examples were also compatible with a telic reading. Here we re-examine the issue, using much stricter criteria and higher-quality data.

The new results strongly support the original conclusion: English atelic path phrases, like telic path phrases, depend on the positive setting of The Compounding Parameter (TCP; Snyder 1995, 2001, 2007, 2011, 2012). We argue that TCP is fundamentally a parameter of the syntax-semantics interface, and that [+TCP] languages have a special semantic operation that can be used for novel compounds, telic path phrases, and also atelic path phrases.

2. Background

Snyder (1995) first observed that certain syntactic structures are found only in languages that allow fully productive (or “creative”) root compounding, illustrated for English in (1).

(1) a. banana box (e.g. 'box for bananas', 'box in the shape of a banana', etc.)

b. faculty lab space committee chair

Here "creative" refers to the creative aspect of human language. For example, native speakers of a given language freely create sentences that may (for all they know) be novel, and they reasonably expect other speakers to understand them. In languages like English and German, the same is true for the creation of bare-root endocentric compounds. Thus, banana box in (1a) can be used as a completely novel term, to refer to a box that has a contextually salient connection to bananas. As long as the listener has enough information to see the connection, use of the term is completely natural and appropriate. In languages that lack creative compounding, such as Spanish, one finds endocentric compounds that are lexicalized, but one cannot create new ones at will.

* The present paper is an off-shoot from the work presented at GLOW-in-Asia 2012 by William Snyder, who is extremely grateful for the questions and comments received there. The authors are likewise grateful for the comments received when an earlier version of this work was presented at the 2011 Workshop on Verbal Elasticity, at the Universitat Autònoma de Barcelona. All remaining errors are of course our own.
Namiki (1994) observes that creativity of compounding, in this sense, is tightly linked to the possibility of recursive compounding, as illustrated for English in (1b): A language has creative compounding if and only if it has recursive compounding. Indeed, in some languages the child's most reliable cue that compounding is a "creative" process in her target language may be adults' use of recursive compounds (Roep, Snyder & Hiramatsu, 2002).

Snyder (1995) argued that languages permitting separable-particle constructions as in (2) or adjectival resultatives as in (3) are also consistently languages with creative compounding.

(2) V-NP-Particle constructions:
   a. lift the box up
   b. turn the light off

(3) Adjectival resultative constructions:
   a. beat the metal flat
   b. wipe the table clean

The basic pattern can be seen by comparing English with Spanish, as in (4-6).

(4) Root compounding
   a. English: frog chair (e.g. 'chair in the shape of a frog')
   b. Spanish: *rana silla (lit. 'frog chair'), *silla rana (lit. 'chair frog')

(5) Verb-NP-Particle constructions
   a. English:  Mary pulled the top off
   b. Spanish:  María tiró el tapón (*de)
                Maria pulled the top off-of

(6) Adjectival resultative constructions
   a. English:  John beat the iron flat
   b. Spanish:  Juan golpeó el hierro (*plano)
                Juan beat the iron flat

Snyder (1995, 2001) provided two different types of evidence to support a connection of particles and resultatives to creative compounding. First he conducted a modest cross-linguistic survey, using native-speaker consultants, with the results summarized in Table 1.1

Note that the implicational relationship in Table 1, at least for particles and compounds, is one-way. That is, the availability of separable particles (e.g. the V-NP-Particle construction of English) entails the availability of creative compounding, but creative compounding does not entail the availability of separable particles (cf. Japanese, American Sign Language).

Snyder (1995, et seq.) proposes to derive the pattern in Table 1 from a grammatical parameter, The Compounding Parameter. Languages with a positive setting of the parameter ([+TCP], e.g. English) allow both creative compounding and at least some (but not necessarily all) of the associated syntactic structures, while the [-TCP] languages (e.g. Spanish) allow neither. In other words, [+TCP] is one of the pre-requisites, but not necessarily the only pre-requisite, for structures like the English V-NP-Particle construction.

---

1 Note that Table 1, drawn from Snyder 2011, includes some updates to earlier versions. In particular, Basque was initially classified as a language with bare-root endocentric compounding like that found in English, but discussion with experts on the language led to the realization that that classification was probably erroneous.
Table 1. Crosslinguistic survey (from Snyder 2011):

<table>
<thead>
<tr>
<th>Language</th>
<th>Separable particles?</th>
<th>Adjectival resultatives?</th>
<th>Creative compounding?</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Austroasiatic)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Khmer</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>(Finno-Ugric)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Estonian</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>(Germanic)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dutch</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>(Sino-Tibetan)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mandarin</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>(Tai)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thai</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Japanese</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>American Sign Language</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Basque</td>
<td>No</td>
<td>No</td>
<td>No?</td>
</tr>
<tr>
<td>(Afroasiatic)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Egyptian Arabic</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>(Austronesian)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Javanese</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>(Romance)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spanish</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>(Slavic)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serbo-Croatian</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

This first type of support for TCP, from a very small-scale cross-linguistic survey, was of a preliminary character, but it was supported by strong converging evidence from a second source, the time course of child language acquisition. Snyder (1995, 2001) examined longitudinal corpora of spontaneous speech from ten children in the CHILDES database (MacWhinney 2000). There he determined the age of acquisition for V-NP-Particle constructions, and for creative endocentric compounding. The age of acquisition was taken as the age of FRU (First clear use, followed soon after by Regular Use; cf. Stromswold 1996).²

The findings were dramatic: Each child began producing novel compounds (i.e. bare-root endocentric compounds that were not lexicalized, had never before been used by the adults in the recordings, and were treated by the interlocutors as novel) at almost exactly the point –

² Note that unlike particles, adjectival resultatives were not examined. The reason was the low frequency of English adjectival resultatives in the speech of both children and adults. Low frequency makes it very difficult to determine an age of acquisition from samples of spontaneous speech.

In Japanese, however, creative compounding is acquired on average about one year later than in English. Sugisaki and Isobe (2001) have tested the prediction of TCP for Japanese adjectival resultatives using experimental tasks (elicited production, EP; and truth-value judgement, TVJ) that can be used appropriately with three-year-olds (but not with two-year-olds). Overwhelmingly, the children who succeeded at their TVJ task for resultatives were the same ones who produced novel N-N compounds in their EP task.
frequently in the very same recording – where he or she began producing V-NP-Particle constructions. These data are presented graphically in Figure 1.

The results of a correlation analysis were highly significant ($r=.98$, $t(8)=12.9$, $p<.0001$), and remained extremely strong even when measures of general language development were controlled for, using a partial correlation procedure. Similar findings were obtained in a follow-up study using corpora for a total of 19 American and British children (Snyder 2007).

![Figure 1: Ages (in years) of FRU of novel compound vs. V-NP-Particle](image)

A major question is what the proposed parameter, TCP, actually does. A number of leading proposals have come from research examining the possible role of TCP in the linguistic expression of motion events (Beck & Snyder 2001a,b; Snyder & Lillo-Martin 2005; Zubizarreta & Oh 2007; Gehrke 2008). Here we will focus on the proposals of Beck and Snyder ('Rule R'), and of Snyder and Lillo-Martin ('Rule C', or 'Generalized Modification').

Building on work of Stechow (1995), Beck & Snyder (2001b) proposed a semantic interpretation of TCP: [+TCP] languages can use Rule R for semantic interpretation, which in turn makes it possible to have resultatives and/or V-NP-particle constructions.

(7) **Rule R:**

If $\alpha = [\gamma s c \beta]$, $\beta'$ is of type $<s,<\tau ,t>>$, and $\gamma'$ is of type $<e,...<e,<s,<\tau ,t>>>$ (where $\gamma'$ is an n-place predicate),

then $\alpha' = \lambda x_1...\lambda x_n \lambda w \lambda t. \text{CAUSE}_{w,t} (\lambda w'\lambda t'. \gamma_{w,t'}(x_1)...(x_n), \lambda w''\lambda t''. \text{BECOME}_{w'',t''}(\beta'))$.

For example, semantic composition by Rule R can combine an activity verb ($\gamma$) and small clause ($\beta$) to form an accomplishment predicate $\alpha$: beat the metal [PRO flat] = beat the metal and thereby CAUSE it to BECOME flat; lift the box [PRO up] = lift the box and thereby CAUSE it to BECOME up.

Crucially, Beck & Snyder (2001a,b) argued that in [+TCP] languages this analysis can be extended to telic path phrases (e.g. walk to the summit = walk and thereby CAUSE oneself to BECOME at the summit). Their account predicts that telic (i.e. 'resultative') path phrases may combine with simple activity verbs to form accomplishment predicates only in [+TCP]
languages. Indeed, Spanish contrasts with English in exactly the way predicted, as illustrated by the following examples from Aske (1989):

(8) a. **English**: John walked to the summit in an hour.
    b. **Spanish**: Juan caminó hasta la cima (*en una hora).

In (8a), English allows the telic path phrase *to the summit* to convert an activity verb *walked*, which is ordinarily incompatible with the modifier *in an hour*, into an accomplishment predicate that can take the modifier. In (8b), Spanish permits the phrase *hasta la cima* to function in a way similar to the English phrase *to the summit* (although *hasta* is probably closer in meaning to 'until' than 'to'). Yet, the predicate *caminó hasta la cima* remains a simple activity, like *caminó*, not an accomplishment. For Beck and Snyder, the conversion of an activity into an accomplishment (using a telic path phrase or any other type of result phrase) is a consequence of applying Rule R, and in [-TCP] languages Rule R is simply unavailable.

The connection made by Beck and Snyder to path phrases drew attention to the fact that TCP overlaps quite a bit with the verb-framed/satellite-framed typology of Talmy (1985, 1991, 2000). Talmy’s typology distinguishes between “satellite-framed” (English-type) languages and “verb-framed” (Spanish-type) languages. In a satellite-framed language, path of motion is normally expressed by a “satellite”— a particle or PP (Talmy 1985:62)— which in (9a,b) leaves the verb free to encode the manner of motion:

(9) a. The rock slid/rolled/bounced down the hill.
    b. The napkin blew off the table.

In a verb-framed language such as Spanish, however, path is normally expressed in the main verb (Talmy 1985:69):

(10) La botella entró a la cueva (flotando).
    ‘The bottle floated into the cave’

According to Talmy (e.g. 1985:68,104), both directional particles and adjectival resultatives are typical of satellite-framed languages. This naturally suggests the possibility of identifying Talmy’s satellite-framed languages with Snyder’s [+TCP] languages.

This move is actually a bit dangerous, as discussed in (Snyder 2012). One reason is that Talmy’s proposals concern language typology, in the sense of the types of surface expression that are typically employed in a given language. In contrast, TCP is a proposal within parameter theory, where the goal is to precisely specify the sentence-meaning pairs that are possible in a given language, in terms of abstract grammatical characteristics.

Nonetheless, the parallels are suggestive, and led Snyder, Felber, Kang & Lillo-Martin (2001) to wonder if Beck and Snyder’s approach might have been too restrictive. Specifically, the path phrases tied to the [+TCP] setting in Beck and Snyder’s proposal were limited to those that specify an end-point location for the Theme (e.g. *walk to the store*), and that are therefore effectively result phrases. Yet, in Talmy’s work, no sharp distinction is made between telic (resultative) and atelic path phrases (e.g. *walk in circles*). Perhaps the distribution of both telic and atelic path phrases is influenced by the setting of TCP, and if so, perhaps Rule R should be replaced (at least in the formulation of TCP) by a more general composition rule, one that can apply to atelic path phrases.

To evaluate these possibilities, Snyder et al. (2001) returned to the evidence from child language acquisition. They found that the age of FRU for a motion verb with a non-resultative (atelic) path phrase was tightly correlated with the FRU of novel compounding (*r*=.91,
Thus, their results indicated that a positive setting of TCP affects atelic path phrases as well as telic path phrases.

In order to account for these findings, Snyder & Lillo-Martin (2005) proposed that the [+TCP] setting makes available a semantic composition rule ('Rule C', or 'Generalized Modification', to be discussed below) that is much more general than Beck & Snyder’s (2001) Rule R. Yet, Gehrke (2008) questions the need for such a general composition rule, on the grounds that some of the data points given as examples of "atelic" path phrases in (Snyder et. al. 2001) actually seem compatible with a telic interpretation, a concern which we address in our new study.

3. The Study

3.1. Predictions

Beck & Snyder (2001), as well as Gehrke (2008), predict that there will be no systematic relationship between atelic path phrases and novel compounding as children acquire a [+TCP] language. Snyder & Lillo-Martin's (2005) account, in contrast, predicts that atelic path phrases (e.g. with pure manner-of-motion verbs) will become grammatically possible for the child at the same time as novel compounding.

3.2. Method

Longitudinal corpora from eleven American English-speaking children were downloaded from CHILDES (MacWhinney 2000) and searched for FRU's of novel compounds and atelic path phrases. Corpora that did not meet the standards outlined in Snyder (2007) were excluded from study. These stringent standards state that the corpora must “follow monolingual English acquiring children without any known developmental abnormality, that are based on audio- or videotape recordings of spontaneous speech (rather than simple diary notes, or recordings of controlled experiments), that cover a time span of at least nine months, that begin when the child is no older than 2;03, that have an average gap between recordings of no more than one month, that contain at least 10,000 child utterances, and that are described in the available documentation as being in a finished state” (Snyder 2007: 53).

Utterances judged to be potential FRUs of atelic path phrases were then evaluated as follows:

(i) Transcribers' notes in the transcripts and original video data (where available) were used to clarify the most likely intended meaning of the child’s verb + path phrase combinations.

(ii) Highly ambiguous path phrases were simply discarded from analysis, in favor of utterances whose physical and/or pragmatic context strongly favored an atelic interpretation.

(iii) Path phrases involving an “un-location” were not counted as atelic. For example in Snyder et al. (2001), expressions like take the record off were treated as atelic, on the grounds that the eventual end-point of the record was left unspecified. Yet, an alternative view (Aske 1989) is that the end-point in this type of example is an "un-location": The record was originally on the turntable, and it became not on the turntable. From this perspective such expressions are better viewed as telic, and were therefore excluded in the present analysis.
3.3. Results

The FRU of a motion verb with an atelic path phrase is given in Table 2 for each of the eleven children studied.

Table 2. FRUs of atelic path phrases with a motion verb

<table>
<thead>
<tr>
<th>Child</th>
<th>Utterance</th>
<th>Onset Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adam</td>
<td>move away</td>
<td>2:03.04</td>
</tr>
<tr>
<td>Eve</td>
<td>Jody coming down the stair</td>
<td>1:09.00</td>
</tr>
<tr>
<td>Lily</td>
<td>mouse is going down the slide</td>
<td>1:10.23</td>
</tr>
<tr>
<td>Naima</td>
<td>run around hay</td>
<td>1:05.11</td>
</tr>
<tr>
<td>Naomi</td>
<td>going up</td>
<td>1:10.29</td>
</tr>
<tr>
<td>Nina</td>
<td>slide down</td>
<td>2:01.15</td>
</tr>
<tr>
<td>Peter</td>
<td>tape go round</td>
<td>2:00.10</td>
</tr>
<tr>
<td>Sarah</td>
<td>climb up the stairs</td>
<td>3:01.03</td>
</tr>
<tr>
<td>Shem</td>
<td>wheel turn around</td>
<td>2:03.02</td>
</tr>
<tr>
<td>Violet</td>
<td>turn around [spinning a toy]</td>
<td>1:09.27</td>
</tr>
<tr>
<td>William</td>
<td>let's go this way</td>
<td>2:04.01</td>
</tr>
</tbody>
</table>

Each of the uses in Table 2 was followed soon after by additional uses, involving different lexical items. In every case the context strongly supported an atelic interpretation. For example, Adam's utterance move away might initially appear to involve an "un-location," as discussed above, but in context the particle away clearly expresses the path of motion, rather than movement out of a specific starting location. (The child is describing a cat that he and his mother had seen in the park. According to the child, the cat first moved farther away from them, and then began to climb up something, possibly a tree.)

The FRU's of atelic path phrases in Table 2 are closely correlated with the FRU's of novel compounding in Table 3 ($r=.9405, r^2=.8833 \ t(9)=8.3, p<.001$). This relationship is first shown graphically in Figure 2. Note that one child, Sarah, is something of an outlier in the sense that she has an unusually late age of FRU for atelic path phrases, even relative to her late age for first novel compounds. This is plausibly due to Sarah having an especially low frequency of atelic path phrases in her spontaneous speech. (See below for similar findings in the children Lily and Naima.) For the sake of comparison we provide a second graph (Figure 3) with Sarah removed. Nonetheless, to be conservative we will rely on the calculations where Sarah is included in the sample, since there was no prior reason to exclude her.
Figure 2. Ages of FRU for novel N-N compounding vs. atelic path phrases

\[ y = 1.1002x - 0.0804 \]
\[ R^2 = .8833 \]

Figure 3. Ages of FRU for novel N-N compounding vs. atelic path phrases, with outlier (Sarah) excluded.

\[ y = 0.8444x + 0.3738 \]
\[ R^2 = .9427 \]

The correlation in Figure 2 remains significant even after mathematically "removing" the portion of the data that one could explain in terms of more general differences across children in their overall rate of linguistic and conceptual development. For purposes of this analysis we calculated two different measures of the children's general development. First, given that children's (and adults') novel compounds are overwhelmingly N-N compounds, the FRU of lexical N-N compounds provides a sensitive index of any delays in novel compounding due to morphophonological difficulty. In other words, to the extent that a child's acquisition of novel compounding might be held back by more general difficulties with the phonology or morphology of the actual surface form of a compound, those difficulties should apply equally to lexical compounds, and should result in a corresponding delay there too.
Second, the FRU of attributive adjectives provides a sensitive index of any delays in novel compounding that are due to general conceptual difficulties with the relationship between a modifier and a head. By mathematically removing all the variability in our data that could be explained in terms of these two measures of more general development, we can see if there remains a strong, statistically significant association (i.e. a significant "partial" correlation) that is specific to true, novel compounding and atelic path phrases.

The results are that the correlation of novel compounding with atelic path phrases remains robustly significant even after we mathematically remove the variability that can be explained by the ages of FRU for lexicalized compounds ($r_{\text{partial}}=.835$, $t(8)=4.3$, $p<.0026$), and even after removing the variability that can be explained by the ages of FRU for simple adjectives modifying nouns ($r_{\text{partial}}=.845$, $t(8)=4.46$, $p=.0021$).

Table 3. Age of acquisition (in years) of atelic path phrases and novel N-N compounds

<table>
<thead>
<tr>
<th>Child</th>
<th>N-N compound</th>
<th>Atelic path phrase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adam</td>
<td>2.26</td>
<td>2.26</td>
</tr>
<tr>
<td>Eve</td>
<td>1.58</td>
<td>1.75</td>
</tr>
<tr>
<td>Lily</td>
<td>1.74</td>
<td>1.90</td>
</tr>
<tr>
<td>Naima</td>
<td>1.38</td>
<td>1.45</td>
</tr>
<tr>
<td>Naomi</td>
<td>1.92</td>
<td>1.91</td>
</tr>
<tr>
<td>Nina</td>
<td>1.96</td>
<td>2.12</td>
</tr>
<tr>
<td>Peter</td>
<td>1.87</td>
<td>2.03</td>
</tr>
<tr>
<td>Sarah</td>
<td>2.59</td>
<td>3.09</td>
</tr>
<tr>
<td>Shem</td>
<td>2.23</td>
<td>2.21</td>
</tr>
<tr>
<td>Violet</td>
<td>1.73</td>
<td>1.82</td>
</tr>
<tr>
<td>William</td>
<td>2.34</td>
<td>2.34</td>
</tr>
</tbody>
</table>

Examining the ages in Table 3 carefully, one will notice that the age of FRU for novel compounding is sometimes almost identical to that of atelic path phrases, but other times is a good bit earlier. (No child, however, begins using atelic path phrases substantially earlier than novel compounds.) This systematic relationship results in a significant ordering effect by paired $t$-test ($t(10) = 2.60$, $p = .0265$).

In principle this could mean that atelic path phrases have a second grammatical pre-requisite, in addition to the parameter-setting of [+TCP] required for novel compounding, and that some children acquire this second pre-requisite substantially later than [+TCP]. Yet, a simpler explanation strikes us as more plausible: Atelic path phrases are used much less often than novel compounds in children's spontaneous speech, even when both are clearly available to the child. Hence, even if compounding and atelic path phrases become grammatically available to the child at precisely the same age, we expect that we will often encounter compounds in a child's speech sooner than atelic path phrases, simply by the luck of the draw.

To assess the plausibility of this frequency-based explanation, we performed detailed case-studies on two children (Lily and Naima) who have a large gap between the two FRU's. The statistical method we applied was a binomial test, based on the relative frequency of the two structures in the child's own speech, after atelic path phrases had begun to appear.

In the case of Lily, the FRU of novel compounding was found at age 1;08.28 (1 year, 8 months, 28 days). The FRU of atelic path phrases was found at age 1;10.23. Before the latter FRU, Lily produced exactly three novel compounds. In a sample of later transcripts, after the latter FRU, Lily produced five novel compounds and only one additional atelic path phrase.
Under the null hypothesis that atelic path phrases were available to Lily just as early as novel compounds, and had the same relative frequency of use as seen in later transcripts, the exact probability of sampling at least three novel compounds simply by chance, before the first occurrence of an atelic path phrase, is \( p = (5/(1+5))^3 = .571 > .05 \). Hence, the observed gap is not statistically significant.

The corresponding probability for Naima is also not significant: \( p = (4/(4+2))^3 = .287 \text{ NS} \).

While it is always difficult to interpret null results, in these cases we take the lack of significance as an indication that the findings are fully consistent with what is expected based simply on the relative frequencies of novel compounding versus atelic path phrases. There is not as yet any strong evidence for a more complex model, involving a second grammatical pre-requisite for atelic path phrases.

4. Discussion

The results of our study are fully consistent with the results of Snyder et. al. (2001), even though we restricted our attention to the highest-quality longitudinal corpora available, and to path phrases that met strict standards for being atelic. Thus, the findings constitute strong acquisitional evidence for a version of The Compounding Parameter that is relevant to both telic and atelic motion predicates. Such an account is laid out by Snyder and Lillo-Martin (2005), who argue that [+TCP] languages have available a highly general semantic composition rule, which they termed 'Rule C'. An updated version of this rule, now termed 'Generalized Modification' (Snyder 2011, 2012), is given in (11).

(11) Generalized Modification (GM):

\[
\text{If } \alpha \text{ and } \beta \text{ are syntactic sisters under the node } \gamma, \text{ where } \alpha \text{ is the head of } \gamma, \text{ and if } \alpha \text{ denotes a kind, then interpret } \gamma \text{ semantically as a subtype of } \alpha's \text{ kind that stands in a pragmatically suitable relation to the denotation of } \beta.
\]


The mechanism of GM can be used to obtain a semantic interpretation for each of the distinctive structures that are found in [+TCP] languages. For example, bare-root compounds can be created quite easily in any language, using the syntactic operation of Merge, but to assign a semantic interpretation to a novel compound we need GM. Thus, the English N-N compound \textit{frog man} in (12) is appropriately used in a wide variety of contexts, with a variety of meanings, and is not restricted to its lexicalized meaning of 'underwater diver'. In the right context, a frog man might also be a man who looks like a frog, keeps frogs as pets, or conducts scientific research on frogs.

(12) Novel compounding:

\[ \text{©} \textit{frog man} \text{©} = \text{man of a type related to frogs} \]

All of these novel, non-lexical uses follow from the denotation provided by GM: a subtype of the semantic kind 'man' that stands in a pragmatically suitable relation to the denotation of \textit{frog}. What is pragmatically suitable will vary, of course, depending on the context of utterance.

The mechanism of GM can also be used to compose a path phrase (whether telic or atelic) with a simple manner-of-motion activity verb. An example is provided in (13). As discussed at some length in (Snyder 2012), GM can apply to individual-kinds, like 'man', or
equally well to event-kinds, such as a "floating" event. Yet, where the human conceptual system can provide countless possible relations between an individual-kind like 'man' and an et-predicate like 'frog', the possibilities for event-kinds are severely restricted. In the case of (13), one of the few relations available between an activity-type event-kind and a locative small clause is the relation between the two subparts of an accomplishment event, the "development" and the "culmination" (in the terms of Parsons 1990). This possibility yields the correct interpretation for (13) when it is understood on a telic reading: The bottle underwent a floating type of motion, beginning at some point that was not under the bridge, and eventually reached an endpoint that was under the bridge.3

(13) Telic path phrases:

\[
\text{[The bottle]$_1$ floated [PRO$_1$ under the bridge]} \]

= There exists a (past) event of the bottle floating, and this event is of the kind associated with the bottle being in a location under the bridge.

= There exists a (past) accomplishment-type event whose development consists of the bottle floating, and whose culmination is the bottle coming to be under the bridge.

Note that (13) can also receive an atelic reading, where under the bridge denotes a path of motion, rather than a resultant location. In the terms of (Gehrke 2008), we might take the PP (at least on this reading) to denote an atemporal sequence of points in space. In this case GM will also succeed in providing a suitable meaning for the VP, as follows: 'There exists a (past) event of the bottle floating, and this event is of the kind associated with a sequence of points passing under the bridge'. We take it that the human conceptual system will ultimately interpret this type of semantic denotation as something along the lines of, 'There exists a (past) event of the bottle floating, and in this event the Theme (i.e. the bottle) moved along a path passing under the bridge'. Admittedly this is just a rough paraphrase, and we will have to leave it to the semantics community to see if the details of this suggestion can be made to work.

Crucially, though, on both the telic (resultant-location) reading of (13) and the atelic (path-of-motion) reading, GM is proposed to play a vital role. In a language that is [-TCP] and therefore lacks GM, we expect that the direct counterpart to (13) will lack both of these readings, and that the PP corresponding to under the bridge will simply be understood as specifying the location of the entire event. In the case of Spanish, at least, this is known to be correct (Talmy 1985).

5. Conclusion

We have now presented evidence that English-learning children acquire the process of novel bare-root endocentric compounding concurrently with atelic path phrases. This evidence lends strong support to a version of The Compounding Parameter like that of Snyder

---

3 Interestingly, Berit Gehrke (p.c.) has noted variation between British versus American speakers of English, in that the British speakers she has consulted resist the telic interpretation of (13), while the American speakers accept it. The British speakers instead favor the atelic-path interpretation discussed below, in which the bottle's path of motion passes under the bridge but continues to points unknown, no longer under the bridge. This variation across speakers could be due to a difference in the precise lexical semantics of float. If so, we expect that the same speakers who resist a telic reading of (13) will nonetheless accept such a reading if a different manner-of-motion verb (e.g. walk) replaces float. We leave this as a topic for future investigation.
& Lillo-Martin (2005): The mechanism in a [+TCP] language that allows creative compounding can also be used for both telic and atelic path phrases.

References

SYNCRETISM DISTRIBUTION MODELING:
ACCIDENTAL HOMOPHONY AS A RANDOM EVENT

Uli Sauerland\(^1\) and Jonathan David Bobaljik\(^2\)
Harvard University/ZAS\(^1\) and University of Connecticut\(^2\)

1. Introduction

The morphological analysis of paradigms generally proposes a distinction between accidental and systematic homophony. No specific assumptions are usually made about the distribution of accidental homophony, though. Therefore current assumptions cannot prove satisfactorily what should be regarded as systematic in morphology. We propose that accidental homophony should be assumed to be a random event in the statistical sense with a constant probability across languages and across paradigms. This approach allows us to assign a likelihood to any actual typological distribution of syncretism given a morphological analysis. And by computing such likelihoods for a range of analyses, we can then apply maximum likelihood analysis to determine the best analyses. Hence, the statistical foundation allows us to empirically test morphological analyses that include accidental syncretism. In this paper, we primarily introduce the conceptual and mathematical foundations of a statistical modeling technique, Syncretism Distribution Modeling, and show how it overcomes the problem of accidental homophony. In addition, we apply the technique to show that person paradigms must involve both accidental homophony and systematic syncretism.

1.1. The Problem of Accidental Homophony

To draw the distinction between accidental homophony and systematic syncretism poses a general problem to linguists of all stripes. There are some clear cases of both types. For example, the English words *bank* ‘financial institution’ and *bank* ‘side of a river’ are a textbook example of homophony: two distinct words (or morphemes) that, due to the vagaries of history, happen to share the same sound. The different meanings are readily seen in translation, where other languages have two different words for these concepts, e.g., in German *Bank* (‘financial inst.’) and *Ufer* (‘shore’). Consider now the English second person pronoun *you*. This single word in English also corresponds to multiple words in other languages, for example, German uses *du* for a singular addressee, and *ihr* for plural addressees. One possible analysis of English would

\(^*\) For helpful comments, we are grateful to Daniel Harbour, Irene Heim, Roni Katzir, Gereon Müller, Rachel Ward, Kazuko Yatsuhiro, and to audiences at GLOW 35 at Potsdam University, GLOW Asia 9 at Mie University, MIT, and the Universities of Connecticut, Iceland, Jerusalem, Leipzig, and Tohoku (Sendai). The research reported on in this work is in part supported by funding from the Humboldt foundation, the German Ministry for Education and Research (BMBF), grant number 01UG0711, and the US National Science Foundation, grant #BCS-0616339.
be to posit two words *you* that share the same sound. Yet rather than positing homophony, it is common to think of English *you* as a case of under-differentiation: English pronominal morphology neutralizes a contrast that is made in other languages, namely, that between singular and plural second persons. The term *syncretism* refers, in the morphological literature, to such under-differentiation in a paradigm, where a single *exponent* (i.e., phonological unit) corresponds systematically to multiple cells. In general, though, there is no theory-independent procedure to determine whether a given case of homophony is due to accidental homophony or systematic syncretism (see also Harbour 2008).

In this paper, we address the distribution of homophony patterns in person paradigms. The English and German personal pronoun paradigms are shown in figure 1. Both languages have three grammatical persons and two numbers, yielding a six-celled *paradigm*. However, whereas English has five distinct (combinations of) sounds for the six cells, German has six. The relation of homophony between cells defines for each language a partition of the cells into equivalence classes. In the following, we use the term *D-Partition* for the partition of the cells of a paradigm defined by homophony. The d-partitions of English and German personal pronouns are indicated by the abstract diagrams below the paradigms in 1. In the following, we are only concerned with the d-partitions, never with the actual phonological content of the cells. Thus, Itelmen, an indigenous language of the Kamchatka peninsula in Russia, shows a 6-way contrast like German. Although the actual phonological forms are entirely unrelated to their German counterparts, we consider Itelmen and German to constitute distinct tokens exemplifying the same d-partition. We note, in addition, that in thinking of paradigm distribution as a partition problem, we ignore the geometric layout of the standard presentation of a paradigm. A paradigm space is an unordered set of feature combinations, and the partitions are simply the possible subsets of this set. There is no meaningful sense in which cells are adjacent to one another or not. We use a constant layout only to facilitate easy identification of correspondences among different partitions.

Our interest in d-partitions in person paradigms is due to the observation that many possible d-partitions are very rare or don’t occur at all, while others are very frequent crosslinguistically (Forchheimer 1953, Cysouw 2003). The number of possible partitions of an *n*-membered set is the Bell number *Bₙ*, which grows exponentially with *n*. For four cells, there are *B₄* = 15 different d-paradigms, any six-celled paradigm has *B₆* = 203 possible d-partitions, and for eight cells there are 4140 possible d-partitions. We discuss in more detail in the following sections evidence that indicates that the frequency distribution of d-partitions across languages requires an explanation. There have already been some attempts to explain this extremely skewed distribution, though these have been for the most part qualitative, rather than quantitative (but see Pertsova 2011). Accounts in the morphological literature posit a universal feature inventory, which allows observed (and in particular common) syncretic patterns (d-partitions) to be described as neutralizations of underlying contrasts (see Bobaljik 2008a, Harley and Ritter 2002, Wechsler 2010). Such approaches therefore argue for language-universal constraints against some types of syncretism. But in attempting such a study, a significant problem presents itself, namely the problem of distinguishing the accidental from the systematic—homophony from syncretism.

As we mentioned above, it is not clear how to draw the distinction between homophony and syncretism in general. Consider one relevant case where different proposals have been made: the German verbal agreement paradigm, also shown in figure 1. Unlike in the German pronouns, in the verb endings there are only 4 phonologically distinct endings spread over the six cells; the first and third person plural endings are identical (*-en*), as are the second person plural and third person singular (*-t*). In principal, both of these two cases of homophony could
be either accidental or syncretism. For example, it may be that the German speaker’s mental grammar includes six exponents for the present tense endings, but that it just so happens that some of the listed endings have the same phonological realization: (1sg = -e, 1pl = -en, ... 3pl = -en, etc.). On this view, the underlying structure of the paradigm doesn’t correspond to the d-partition, but is the same as that of German and Itelmen pronouns. In what follows, we will use the term **M-Partition**, to refer to the underlying, abstract grammatical structure of the paradigm. Only if there is no accidental homophony, the m-partition is the same as the d-partition. In the analysis just discussed, the m-partition is the same as the paradigm space, that is, the maximal partition of that space. A second conceivable m-partition for German verbal agreement is the d-partition. In addition, two intermediate m-partition exist in this case. All four possible m-partitions in this case are shown in figure 2.1 We compare partitions in terms of their **coarseness**. Partition \( m_1 \) is at least as coarse as \( m_2 \) if any two members of the same equivalence class of \( m_2 \) also belong to the same equivalence class in \( m_1 \). Coarseness partially orders the \( m \)-partitions in figure 2, with the coarsest \( m \)-partition on the right being coarser than the other three, and the finest on the left finer than the other three. But the middle two \( m \)-partitions in figure 2 are not ordered relative to each other by coarseness.

The decision of which \( m \)-partition is correct in a case like German is usually made on the basis of theory internal criteria. As we mentioned, at least two morphological analyses for German verb agreement have been offered in the literature corresponding to the two \( m \)-partitions on the right in figure 2. The two analyses of the German facts, cast as ordered rules of exponence (see Anderson 1992, Stump 2001), are given in (1a) and (1b). Analysis (1a) is an instance of the third \( m \)-partition of figure 2, while (1b) instantiates the coarsest \( m \)-partition.

---

1 In general, the number of possible \( m \)-partitions for a given \( d \)-partition is the product \( \prod E e B_{c(e)} \) where \( e \) ranges over the exponents of the \( d \)-partition and \( c(e) \) is the number of cells where exponent \( e \) occurs.
The two analysis also use a different feature logic. Analysis (1a) only uses positively specified features, while analysis (1b) allows appeal to both negatively and positively specified features to admit [-2] as a feature that includes 1st and 3rd person. For this reason, the analysis can be successful without appeal to accidental homophony.

\[
\begin{align*}
(1) & \quad a. & \quad 2 \text{ pl} & \rightarrow -t \\
& & \quad \text{pl} & \rightarrow -en \\
& & \quad 2 & \rightarrow -st \\
& & \quad 3 & \rightarrow -t \\
& & \text{default} & \rightarrow \emptyset \\
& \quad b. & \quad [-2, +\text{pl}] & \rightarrow -en \\
& & \quad [+2, -\text{pl}] & \rightarrow -st \\
& & \quad [+1, -\text{pl}] & \rightarrow \emptyset \\
& & \quad [-1] & \rightarrow -t
\end{align*}
\]

We have kept above to a simplified version of the general problem. As paradigm spaces (and the features involved) expand, the problem of distinguishing the accidental from the systematic grows quickly. Homophony happens. It is thus impossible to know a priori what the correct m-partition is for any given d-partition. In the simple case of the German present tense that we have just looked at, the facts are clear, but we have given three analyses invoking 4, 5, or 6-membered m-partitions. This case also makes clear that the spectre of circularity looms large here, as we seem to be trading a more complex analysis for a reduction of accidental homophony.

There is no agreement in the field as to which is the correct m-partition and how to decide such a question. As Harbour (2008) points out, linguistics has not yet come to grips with the problem of accidental homophony. In analyzing identity of form within paradigms, some authors even deny that there is any internal feature decomposition at all, and thus effectively treat all formal identity as a form of homophony (Cysouw 2003). In some sense, the other extreme position is advocated by Halle and Marantz (2008) who seek to avoid accidental homophony and maximize generalization. Concretely they propose to count across languages the number of cases of accidental homophony a morphological analysis requires minimally, and always choose the analysis with fewer instances of accidental homophony. This proposal however is problematic for at least three reasons: For one, the analyses compared may differ substantially in their internal complexity and it would not seem reasonable to adopt a much more complex analysis (for example, one with negative feature values in addition to positive ones) only to avoid a single instance of accidental homophony. Secondly, the choice of analysis could depend very strongly on the languages for which data is available. The oversight of a single language that would require accidental homophony may lead us to prefer a wrong analysis.
Finally, there might be cases where one analysis that requires least homophony requires always homophony between the same two cells. A different analysis, however, leads to more instances of homophony overall, but these are distributed evenly over different pairs of cells. In such a case, we would intuitively prefer the latter analysis, but Halle and Marantz (2008) would have us advocate the former. The problem of accidental homophony also cannot be addressed in the same way as some other problems in morphology. One difficulty is that paradigms are typically not just finite, but relatively small and closed. To test the productivity of other morphological processes, researchers have relied on variants of the *wug*-test of Berko (1958): productive processes are extended by speakers to novel or nonce forms they have never heard before, and thus could not have learned. The *wug*-test works fine to demonstrate that past tense is productive, but doesn’t help with our problem: When it comes to person, there is no reasonable way to simply add a new feature or grammatical person to the system.

In the rest of this paper, we propose a new method to evaluate morphological models. The underlying intuition is that accidental homophony can be viewed as a random event. This gives us a lot of predictions, because a random event is a statistical concept: A random event is very likely an even distribution of accidental homophony and very unlikely to have an uneven distribution. For a morphological model and a fixed set of observations, we can then distinguish between the homophonies that can be analyzed as systemic syncretism in the morphological model and the homophonies that must be accidental. Then we can test whether the distribution of the accidental syncretism is likely or unlikely given our assumption that it is a random event. By this method, we can for different models of morphology compute how likely they predict a set of actual observations to be. Since we assume that the real world that the observations are taken from is fully explained by the model and the random variable of accidental homophony, we would then prefer models that result in a comparatively high likelihood for the set of observations we are looking at. In this way, we can select from a number of morphological models the one that assigns the highest likelihood to the real world. In a very loose sense, what we propose is akin to the following exercise: imagine a situation in which a black box contains an unknown number of regular 6-sided dice, that can be thrown, with the total of each throw reported. What we observe directly is some distribution of totals per throw: throw one, total 8; throw two, total 6, throw three, total 7, throw four, total 7. While this observation is consistent with the model where there are six dice in the box, the model with just two dice is clearly the one that best fits the observations. The problem of finding a good morphological model is more complex, of course, but having the dice-detecting analogy in mind may help to think about what we are doing below.

1.2. The Cysouw Data-Set

Our analysis in this paper makes use of the substantial data set available for the domain of person paradigms from a large, cross-linguistic sample originally collected and analyzed by Cysouw (2001, 2003). We now first introduce this data set, before returning to the issue of accidental homophony and the theoretical tools we will make use of. As far as we know Cysouw’s data-set is the largest collection available, and although there are limitations, it has several good properties. Cysouw’s data-set is primarily gathered from existing grammatical descriptions, established by many others. However, the coding of person is well-described, with broad consensus on the observed categories. So, in our view, Cysouw’s data are of high reliability, though that is most likely also the case for other datasets in this area. A more significant advantage is that Cysouw, as already mentioned, arrives at conclusions that are quite different from the ones we will argue for: namely, Cysouw argues that homophony is not indicative of an underlying feature struc-
tured (in common understanding, it is synchronically accidental), while we will show here that such an analysis is inconsistent with Cysouw’s data.² So, we can be sure that the dataset is not selected with a bias towards our conclusion. For these two reasons and also mere convenience, we chose to use Cysouw’s data-set.

Cysouw’s data set consists of a survey of some 265 tokens of paradigms (i.e., d-partitions in our terms) of person marking from the world’s languages, representing 61 types, and indicating the relative frequency of each type in the sample.³ The dataset includes both independent pronouns and inflectional elements (bound affixes). It is by far the most comprehensive survey of its sort, however words of caution are in order regarding interpretation of the quantitative data. Of particular relevance is a skew in favour of the over-representation of rare paradigms. Thus: “[e]very single example of a rare paradigmatic structure has been included in the sample to show the inherent variability of human language ... On the other hand, not every case of the commonly occurring paradigmatic structures is included. After a representative group of a specific common paradigmatic structure has been described, [Cysouw] stopped the collection of more exemplars of such common patterns.” (Cysouw 2003, 22-23). In addition, to control (in part) for the likelihood that shared paradigm structures among related languages arise from common ancestry, rather than from general properties of the person system, Cysouw limited the counts of any individual d-partition to no more than 4 from a given family (Cysouw 2001, 341). Acknowledging that this move limits our confidence in the results we may obtain, we press forward with this data set in any event, as a more accurate sample will have only a more pronounced divide between the common and the rare, and should thus produce cleaner results. Other concerns about the data set concern independence of the data points and are identified below.

Cysouw codes, descriptively, for an 8-cell parameter space of person and number, rather than the six-cell space familiar to speakers of Western European languages and presented with reference to German above. The underlying person categories are speaker/author [1], hearer/addressee [2], the combination of the two [1+2] (“first person inclusive”), and non-participant [3], and these may be singular (minimal) or non-singular (a group containing the designated referent. A language with the full 8-way contrast is Ilocano as shown in figure 3, where also English and German as eight-cell paradigms are shown for comparison.

The minimal inclusive cell has been the subject of much discussion since it was first identified (Thomas 1955). Minimal inclusive pronouns refer to one speaker and one hearer (you and me) and no others, and form a number pair with a corresponding plural (a group containing this duo). These pronouns are not grammatically duals (see Cysouw 2003, Bobaljik 2008a for reviews of the literature).

This 8-cell paradigm space marks the maximal set of person contrasts in the world’s languages. There is never a person-contrast of more than 4 persons in any one grammatical number, and moreover, if there is a four-way contrast, it is always the one in figure 3. Additional contrasts involve cross-classification with independent features such as gender, case, etc.⁴ All known paradigms (d-partitions) can be derived from this as neutralizations of one or more

² Note that Cysouw (2011) can be read as retreating somewhat from his earlier view, acknowledging to some degree the independence of number from person.
³ We rely primarily on Cysouw (2001), since quantitative information is provided in the appendix to that work. Cysouw (2003) offers a few minor corrections and reclassifications of some of the data. At this preliminary stage, our main purpose is to test the model over an existing data set; in ongoing work, we are collecting a more careful sample, allowing for more robust results than Cysouw’s.
⁴ This is by know means a logical necessity; there have been proposals for additional contrasts, famously between multiple hearers [2pl] versus hearer and others [2+3,pl]; yet no example of such putative con-
contrasts. For example, English and German lack an inclusive-exclusive contrast, thus three of the 8 cells have an identical form (e.g., English we).

We note off the bat that Cysouw’s data set, though constructed with an eye to maximizing the representation of rare types, includes only 61 d-partition types out of 4,140 (= \(B_8\)) logically possible types.

Cysouw also presents separately the results for a smaller space, the top four cells of the eight in figure 3, which correspond to the various first-person markers (inclusive vs. exclusive \(\times\) singular/minimal vs. plural). This is a smaller set of possibilities: there are only 15 partitions of a 4-membered set—restricting investigation to this smaller paradigm space will yield more tractable computations in what follows. We show this data in figure 4. Cysouw presents specific numbers for the occurrences of each of the rare patterns, but, as noted above, capped his collection of examples of the common patterns and does not provide specific counts. We have used 15 as a proxy where Cysouw speaks of “dozens”, consistent with Cysouw’s methodology of low-balling the counts of the exceedingly common patterns and artificially amplifying the relative frequency of rare paradigm types. With this restriction, we retain the data from 83 paradigms. The distribution is still an unusual distribution: 5 d-partitions have 15 occurrences each, while another 5 d-partitions don’t occur at all, as Cysouw already notes.

2. Accidental Homophony as a Random Event

In this section, we introduce our main assumption: that accidental morphology is a random event in the statistical sense. Furthermore we show how our main assumption can be used in morphological analysis, specifically by assuming that the probability of accidental homophony is constant across a paradigm and across languages.

2.1. The Rate of Accidental Homophony

Now, consider again accidental homophony. Our proposal is that accidental homophony should be treated as a random event in the comparative typology of paradigms. The major motivation for this proposal is that it captures the intuitions traditionally connected with the notion of accidental homophony (Cysouw 2003, Simon 2005, Bobaljik 2008a).
of accidental homophony; namely, the intuition that through an accidental development in language history—for example, a sound change or the introduction of a loan word—two distinct morphemes end up sounding the same, even though the grammar doesn’t allow a meaningful identification of the two morphemes. The factors bringing about accidental homophony are therefore regarded as historical accidents independent of the paradigm structures examined. The assumption that accidental homophony is random doesn’t entail that it could not be caused. Also random noise in scientific experiments is usually taken to have some cause: In scientific experiments, we assume that besides some variables that are controlled for a wealth of known and unknown additional factors affect the result. Therefore, the measurement may differ even if the variables that are controlled for are constant. But as long as it is justified to assume that the other causes are randomly interfering with the ones we control for, we can apply statistical computations to say something about the effect of the controlled variables. We propose that, in the same way, our determination of paradigm structure is affected randomly by accidental homophony.

Our proposal that accidental homophony is a random event makes it possible to approach accidental homophony with statistical methods. Statistical methods rely on the insight that, while a single observation of a random event is not informative, repeated observation usually is: the cumulated outcomes generally show a highly regular behavior. For this behavior to emerge, the different observations must be independent of each other in the sense of probability theory. In essence this means that all the prior observations must not affect the next one like a dice roll doesn’t depend on previous rolls.

To apply statistical concepts to accidental homophony, we assume that the basic event is the following choice: Do two distinct morphemes A and B accidentally sound the same or not. We assume that accidental homophony in this case has probability $h$, the overall rate of
accidental homophony. If the value of \( h \) was 0.1, the two randomly chosen morphemes (in a paradigm) A and B would have a 1 in 10 chance of sounding the same and a 9 in 10 chance of sounding different. Only if \( h \) was 0, would accidental homophony be completely impossible.

### 2.2. Two Independence Assumptions

We make two further assumptions as working hypotheses: (i) independence across the paradigm, and (ii) independence across languages. These assumptions mean that for any two morphemes of a particular language from a specific paradigm, \( h \) is the likelihood of the two morphemes sounding the same. For the example of person morphology, consider first languages that have an m-partition with four first person morphemes 1, 1+2, 1+3, and 1+2+3. Then \( h \) is the chance of 1+2+3 sounding the same as 1 as well as that of 1+2 and 1 sounding the same, and so on for any other selection of two morphemes from the m-partition. Compare this now with languages that have an m-partition with only two morphemes, 1 and [1+2, 1+3, 1+2+3]. Then the same value \( h \) is the likelihood of these two morphemes sounding the same.

Both independence assumptions may turn out to be false after investigation, but they are the right points of departure. In the absence of evidence showing an quantifiable effect of other factors, the independence assumptions should be our null assumptions. If at some point in time the effect of our other factors on \( h \) can be quantified, the techniques we develop here could be extended to accommodate this in the model. For now, we think both assumptions can be maintained, though they are to some extent idealizations. Consider first independence across languages. Though languages as a whole are certainly not independent of one another, there is to our knowledge no evidence that the distribution of accidental homophony is similar across distinct, related languages. Still care must be taken in the selection of a language sample to ensure that independence of accidental homophony across languages is a defensible assumption, but the Cysouw dataset that we use satisfies these criteria. A second way in which independence across languages is an idealization is that the likelihood of accidental homophony may depend on the phonological inventory of a particular language. Intuitively it seems plausible that a language with a small inventory of possible syllables should have a greater chance of accidental homophony than a language with a large inventory. However, we are not presently aware of any evidence actually showing such an influence, and therefore feel that it would not be justified to include phonological factors in our statistical model at this point. For independence across the paradigm, it is possible that for example, morpheme frequency affects the likelihood of homophony (William Snyder, p.c.). But, in this case too it isn’t sufficient to only look at d-partitions and it therefore is difficult to actually be sure of the plausible effect. So at this point, we resist the inclusion of additional free parameters unless their influence has actually been empirically demonstrated.

### 3. Accidental Homophony and Morphological Models

#### 3.1. Morphological Parameter Spaces

In the following pages, we present a statistical technique to develop the independence assumptions into a tool to evaluate morphological theories against typological frequency data. We will use the name Syncretism Distribution Modeling (SDM) for this technique. We will use SDM to compare different theories of d-partition distributions. We conceive of an analysis of the paradigm patterns as a condition specifying which m-partitions are available as possible parameter settings. For example, one theory of grammar may say that only d-partitions 1, 3, 15 are
available as m-partitions. The set of m-partitions available constitutes a *Morphological Parameter Space*. If d-partitions other than those occurring in the parameter space occur, they must arise via accidental homophony. Note that the d-partition with as many morphemes as cells cannot arise via accidental homophony, so if it occurs, the morphological parameter space must contain this paradigm.

Generally when we have 4 cells, any subset of the set of 15 paradigms would constitute a parameter space. But since four different exponents for first person actually occur, only parameter spaces that contain paradigm 1 are possible. Then, the total number of possible parameter spaces is then $2^{14} = 16192$.

The two parameter spaces that are most important for us in this paper, are the parameter spaces that make the weakest assumptions about universal grammar: Analysis *All-Parameters*, the parameter space that contains all 15 possible m-partitions, and Analysis *All-Accidental*, the parameter space that contains only m-partition 1. Analysis All-Parameters says that all 15 patterns are available as parametric choices. So this analysis assumes that languages can freely vary as to which set of concepts they verbalize in the person paradigm. As far as a universal theory of language, it assumes that part of the description of any language is a set of concepts it verbalizes. Analysis All-Parameters can therefore account for any syncretism as systematic homophony, but it still allows accidental homophony to occur as well. At the opposite extreme, Analysis All-Accidental assumes that only the paradigm with four distinct cells is available as a parametric choice – as explained above, at least this parametric choice must be available because otherwise Ilocano would not be a possible language. The All-Accidental analysis assume that underlying all languages are like Ilocano and that all syncretism is accidental homophony. As far as universal grammar is concerned, the assumptions of the All-Accidental analysis are even weaker than the All-Parameter analysis. Namely, the All-Accidental analysis assumes that the semantic categories are given independent of language and learning a language only requires acquiring the sounds for each of the semantic categories. In a sense, the All-Parameters and All-Accidental analyses are opposites, but both constitute reasonable null hypotheses. Most active morphologists reject both analyses. But, as we discussed in Section 1.1 the argument against either of the two have rested on problematic assumptions about accidental homophony. The main goal of the present paper is to present an new statistical argument against both null hypotheses. Namely, we argue that both null hypotheses fall short of providing an explanation of the cross-linguistic frequencies of paradigm patterns because they predict the actual pattern frequencies to have an extremely low probability. This reasoning is based on precise probabilistic assumptions about accidental homophony that make it possible to compute the likelihoods of the actual pattern frequencies for different analyses.

### 3.2. The Probability of Patterns

In this section, we introduce how to compute, for an m-partition, the likelihood of d-partitions that can arise from it given a rate of accidental homophony $h$. The rate $h$ as we introduced above refers to the likelihood of homophony of two distinct morphemes of an m-partition. The rate $h$ therefore determines for any underlying m-partition $m$ (with #m morphemes) how likely any d-partition $e$ that is at least as coarse as $m$ is to arise. The simplest case is a two-morpheme paradigm. In this case, the application of $h$ is straightforward: the likelihood of the d-partition with just one exponent is the likelihood of the two morphemes being homophonous, i.e. $h$. By complementarity, the d-partition with two non-homophonous morphemes has likelihood $1 - h$.

Now consider paradigms with more than two morphemes. The likelihood of a particular exponent structure arises from a repeated random event, deciding for each morpheme which
other morpheme it is homophonous with if it is accidentally homophonous with another morpheme at all. In probability theory such repeated random events are often conceptualized as drawing balls of different colors from an urn. An important factor is whether the ball is returned to the urn or not after any drawing of a ball. If the ball is not returned, the likelihoods in the second drawing depend on the color of the initial ball, and at some point we would run out of balls of one color altogether. For these reasons, the model without return doesn’t fit our assumptions: e.g. if cells 1 and 2 are already homophonous, the likelihood of cell 3 to be homophonous with 1 and 2 is still $h$ by our assumption that $h$ is a constant for any two cells. Also a model without return of the balls predicts that at some point a particular color is used up. But that would predict that there must be a specific number $n$ such any exponent can only be ambiguous in at most $n$ many ways. Such a numerical limit on homophony is implausible for morphology. In sum we conclude that the random event we are looking at is appropriately viewed as an urn model with return of the balls. In such a model, the likelihood of drawing a ball of the same color stays $h$ even after multiple drawings. There is one further caveat: The urn model initially seems to only allow fractions likelihoods: if the urn contains $n$ balls in total, and $m$ of them are of a specific color, then the likelihood of drawing a ball of that color is $m/n$. We can, however, assume $h$ to vary without a restriction to fractions since any value of $h$ can be approximated to any degree of precision by a fraction.

With these assumptions in place, it is now clear how we can calculate the likelihood of a particular paradigm pattern for a paradigm with $m$-many morphemes. The general formula for a paradigm with $m$ morphemes and $e$ distinct exponents is given in (2) (assuming $0 \leq h \leq 1/(m-1)$):

$$\Lambda_h^m(e) = h^{m-e} \prod_{i=1}^{e-1} (1 - ih)$$

In the case of $m = 2$, the formula in (2) says that the d-partition with $e = 2$ distinct exponents has likelihood $1 - h$ and the one with only one exponent ($e = 1$) has likelihood $s$. If $m = 3$, we can build upon the case of two cells. If cell 1 and cell 2 are homophonous, the chance of cell 3 also being homophonous is also $h$. So overall the chance of accidental homophony of all three cells is $h^2$, while the chance of homophony of 1 and 2, but excluding 3 is $h(1 - h)$. Now consider the case that cell 1 and cell 2 aren’t homophonous. Cell 3 then has chance $h$ of being homophonous to cell 1, chance $h$ of being homophonous to cell 2 and $1 - 2h$ chance of not being homophonous to cell 1 or 2.

Formula (2) entails that the likelihood of two paradigm patterns is the same if they involve the same number of exponents, regardless of the number of times each exponent occurs. We want to show by means of an example that this is correct: With four cells, there are two different possibilities to fill the paradigm with two exponents, $E_1$ and $E_2$: either $E_1$ occurs three times and $E_2$ only once, or $E_1$ and $E_2$ both occur twice. Concretely consider the likelihood of the two patterns $\langle E_1, E_1, E_1, E_2 \rangle$ and $\langle E_1, E_1, E_2, E_2 \rangle$. In the first case, morpheme 2 and morpheme 3 both have likelihood $h$ to be identical to morpheme 1, and morpheme 4 has likelihood $1 - h$. In total then the first pattern has likelihood $h^2(1 - h)$. In the second pattern, the likelihood of morpheme 2 being identical to morpheme 1 is $h$, and the likelihood of morpheme 4 and morpheme 3 being identical is also $h$. Furthermore the likelihood of morpheme 1 and morpheme 3 being distinct is $1 - h$, so we also arrive at $h^2(1 - h)$ as the likelihood for the second pattern.

Finally, consider our assumption that $h$ be no greater than $1/(m-1)$ that we already mentioned above. The reason for this assumption is that for most $h$ greater than $1/(m-1)$, our basic assumption that $h$ is constant for any selection of two distinct morphemes from an
m-partition leads to an inconsistency. Consider the case of 4 morphemes and \( h = 40\% \). Now assume that, in one specific language, morphemes 1, 2, and 3 are all different. The likelihood of morpheme 4 to then be accidentally homophonous to any of 1, 2, and 3 must be equal to \( h \). But, these three likelihoods of mutually exclusive events add up to 120\%. This contradiction shows that \( h \) cannot be as big as 40\%.\(^5\) Note in addition that the formula as stated above would yield negative values in some of the cases ruled out.

### 3.3. Calculating the Best Fit Syncretism Rate for Model 1

In this section, we show how to find \( h \) so that it best fits a set of data in the case of a model where just one parametric choice exists, i.e. what we termed above All-Accidental models. For the case of the four first person cells, we call this model Model 1. The approach developed with just one parametric choice can then be extended for models with more parametric choices by adding up the prediction for each parameter choice. The kind of predictions our approach makes are statistical unless \( h = 0 \). This means that generally any frequency distribution for the \( \text{d-partitions} \) is possible, but crucially the likelihoods of different distributions will differ. The general technique we will then apply to compute a value for \( h \) is maximum likelihood modeling (see also section 3.6).

So our immediate goal is to find the value of \( h \) such that it makes the observed distribution maximally likely for the given parameter space. Concretely, we consider Model 1 containing a single \( m \)-partition. In other words, we assume that the morpheme structure of all 83 paradigms is the one of Ilocano: four different morphemes. For any value of \( h \), the expected frequency of \( \text{d-partition} \ e \) is given by \( 83 \Lambda_h^{\text{d}}(\#e) \) where \( \#e \) is the number of exponents of \( \text{d-partition} \ e \). Consider for example \( h = 0.1 \): The relevant coefficients are \( \Lambda_h^{\text{d}}(1) = h^3 = 0.1\% \), \( \Lambda_h^{\text{d}}(2) = h^2(1 - h) = 0.9\% \), \( \Lambda_h^{\text{d}}(3) = h(1 - h)(1 - 2h) = 7.2\% \), and \( \Lambda_h^{\text{d}}(3) = (1 - h)(1 - 2h)(1 - 3h) = 50.4\% \). From these we can compute the expected distribution when there are 83 total occurrences: 0.083 occurrences are expected for \( \text{d-partition} \) 15 with one exponent, where actually 15 are observed. We also expect 0.75 occurrences each for \( \text{d-partitions} \) 8 through 14 with two exponents each and 6.0 occurrences for each of the three-exponent \( \text{d-partitions} \) 2 through 7. Finally, 41.8 occurrences of the single four-exponent \( \text{d-partition} \) 1 are expected. The expected frequency distribution is, of course, not predicted as the actual frequency distribution, but rather the mean point of a probability distribution. I.e. if we repeatedly selected 83 languages at random and took the mean number of occurrences across these trials, the means ought to converge against this number.

The \( h \) with the best fit between the expected distribution and the observed distribution is determined by the distance between the two. Both the expected distribution and the observed distribution are vectors of 15 values that define points in a 15 dimensional space. One way to do this would be use the Euclidean distance of the two points, i.e. the square root of \( \sum_{\text{e}}(O_{\text{e}} - E_{\text{e}}^{h})^2 \), where \( E_{\text{e}}^{h} \) are the expected (depending on \( h \)) and \( O_{\text{e}} \) the observed frequency of \( \text{d-partition} \ e \). But using Euclidean distances would be sensitive only to absolute differences, whereas intuitively relative differences are also important. Consider, for example, widely different relative deviations between \( \text{E} \) and \( \text{O} \): assume in one component \( E_{\text{e}} = 1 \) and \( O_{\text{e}} = 50 \) while in the other \( E_{\text{e}} = 50 \) and \( O_{\text{e}} = 100 \). With Euclidean distances the latter difference would still count as worse than the former, even though in the former \( O_{\text{e}} \) is off by factor 50. Therefore the use of Euclidean distances would lead us to have to accept much larger relative deviations between \( \text{E} \) and \( \text{O} \) when \( \text{O} \) and \( \text{E} \) are both small than when they are large. This outcome is generally regarded as

\(^5\) For \( h = 1/n \) for any integer \( n \), no problem arises even if \( h > 1/(m - 1) \).
undesirable. Instead we will use the following measure of distance: \( \text{fit}(h) = \sum (O_e - E^h_e)^2 / E^h_e \). While we can not fully justify this decision at this point of the paper, note that this is the kind of measure that the well-known chi-square statistic makes use of. In section 3.6, it will become clear that this distance measure ensures the \( h \) is chosen so that it maximizes the probability of an event of the type of \( O \). The solid line in figure 5 shows fit between expected and the observed distribution for \( h \) within the interval \([0.15, 0.33]\). We used the function `optimize` of the statistics software R to find the minimum of fit automatically: The accidental homophony rate that gives the closest fit is 28.55%.

![Figure 5: Dependence of fit on the rate of accidental homophony for Models 1 (solid line), 2 (dashed), and 2' (dotted).](image)

In sum, we have shown how different rates of accidental homophony lead us to expect a different frequencies of the d-partitions even when all languages share the same m-partition. We can then compute the fit between the expected distribution and an observed distribution and use this to choose the accidental homophony rate \( h \) so that the fit is best. However, in this section we only considered a model where all languages have the same m-partition. The following section shows how we can compute the expected distribution and again find the \( h \) with the best fit.

3.4. Parameter Choice and Model 2

Linguists have used the term parameter to describe linguistic variation (Chomsky and Lasnik 1993). A parameter space captures the space of possible languages, and each language is associated with a parameter choice. In Model 1 there was no meaningful parameter choice – all languages shared the same m-partition. We now introduce Model 2 that allows a parameter choice: Namely it allows a language to be specified for either m-partition 1 (Ilocano) or m-partition 13 (English). Then, the expected distribution \( E^h \) depends not only on the likelihood of accidental homophony, but also on the likelihoods of the two parameter choices. The

---

6 Other statistical procedures rely on other computations of distance. For example, the G-test relies on log-likelihood ratios. We opt for the chi-square based computation at this point because the statistics software R we use has a fuller implementation of the chi-square test, while we will need to implement a Monte Carlo versions of other tests ourselves for the computations in Section 3.6.

7 Or possible a likelihood profile for the possible parameter choices. (Yang 2002)
likelihoods of parameter choices are difficult to determine \textit{a priori} at this point,\(^8\) so we limit our attention to two approaches: On the one hand, an equal distribution (\textit{Equi-Parametrization}) where both m-partitions have likelihood 50\%. And on the other hand, a best fit distribution (\textit{Best Fit Parametrization}) where the likelihood of the m-partitions is optimized to result in the best to an observed distribution. Specifically, we introduce \textit{Model 2} as an equi-parametrization model in the first person case and \textit{Model 2'} as a best-fit parametrization version of model 2. In this subsection, we focus on the equi-parametrization and model 2, and return to the best-fit parameterization and model 2' in the next section.

The equal distribution generally assumes that each parameter choice has the same likelihood. For example, if 5 m-partitions are available as parameter choices, each is assumed to have likelihood 20\%. For Model 2, however, each of the two m-partitions has likelihood 50\%. In the absence of considerations arguing against equal likelihoods, we think the equal distribution is the best starting point. Furthermore in at least one domain of parametrization, an equal distribution is attested: Gilligan (1987) provides data on agreement from a biased sample of 100 languages. He considers which arguments a verb agrees with. There are four parametric choices in this case: no agreement at all, agreement with the subject, agreement with subject and direct object, and agreement with subject, direct object and indirect object. Gilligan’s findings are reproduced in Table 1.\(^9\) The G-test (Sokal and Rohlf 1995) shows that Gilligan’s distribution is not significantly different from an equal distribution (G(3) = 2.1785, p-value = 0.5362).

\begin{table}
\centering
\begin{tabular}{|c|c|c|c|c|}
\hline
parameterization & no Agreement & S only & S and DO & S, IO and DO \\
\hline
\# of languages & 23 & 21 & 31 & 25 \\
\hline
\end{tabular}
\caption{Distribution of agreement in different languages.}
\end{table}

Once the likelihoods of the parameterizations are determined, we can compute the overall expected distribution by computing the expected distribution for each possible parameter choice given the percentage of language with that parameter choice and then adding up the results from the individual parameter choices. In the case of model 2, we show below that this is straightforward. In the general case, assume that the parameter space contains the m-partitions \(m_2, \ldots, m_n\) in addition to \(m_1\), the complete m-partition space which is needed for the analysis of languages without syncretism. We again use \# for the number of morphemes in m-partition \(m\). Note that \#\(m_1 = n\), but \#\(m_i < n\) for \(i > 1\). Assume that the vector \(\vec{a}\) contains the likelihood for each parametrization, i.e. \(a_i\) be the likelihood that a language will have parametrization \(m_i\) for each \(i\). Since \(m_1\) through \(m_n\) are by assumption all the available parameterizations, \(\sum a_i = 1\). In the case of the equi-parametrization, \(a_i = 1/n\) for any \(i \in \{1, \ldots, n\}\).

If \(h = 0\), the expected distribution \(E^h\) equals \(\ell \vec{a}\), where \(\ell\) is the number of paradigms sampled. But if \(h > 0\), the distribution of d-partitions of the languages with different m-partitions can be affected differently depending on which d-partitions are coarser than the m-partition and on the number of morphemes in an m-partition. The former effect is due to the fact that from each m-partition only coarser d-partitions can arise via accidental homophony. The later is the observation that, at the same general rate \(h\), m-partitions with more morphemes are more likely to exhibit accidental homophony than m-partitions with fewer morphemes. Consider for example Model 2 with \(h = 0.1\): Since parameter choice 2 has only two morphemes while choice 1 has four, the expected rate of d-partitions without accidental homophony is 90\%.

\[^8\text{Future work may be able to determine a priori rates, for example by the application of game-theoretic methods.}\]

\[^9\text{Gilligan actually reports the language Waskia as having agreement with the subject and the indirect object. However, Bobaljik (2008b) argues that Waskia actually has a form of suppletion with give, not indirect object agreement. Hence we count Waskia as a language with only subject agreement.}\]
of the language with parameter choice 2, but only 50.4% of the languages with choice 1. Let \( d \) be a d-partitions. If \( d \) is at least as coarse as the m-partition \( m \) of a parameter choice, the \( \Lambda_h \) coefficient introduced above gives the likelihood of d-partition \( d \). As argued above, the coefficient actually depends only on the number of morphemes in \( m \), \#\( m \), and the number of exponents in \( d \), \#\( d \). Furthermore, the likelihood of arising from \( m \) is 0 regardless of \( h \) for any d-partition other than those at least as coarse as \( m \). Therefore, we can define a matrix \( M(h) \) that specifies for each m-partition of the parameters space \( m \) and for each d-partition \( d \), the likelihood of \( d \) arising for a language with parametrization \( m \):

\[
E_{i,j}^h = M_{d_i,m_j}^h = \begin{cases} 
\Lambda_h^{\#m_j}(\#d_i) & \text{if } d_i \text{ is at least as coarse as } m_j \\
0 & \text{otherwise}
\end{cases}
\] (4)

The number of rows of matrix \( M(h) \) is \( n \), the number of m-partitions in the parameter space. The number of columns is the number of possible d-partitions, i.e. \( B_c \), the Bell number for the number of cells in the paradigm under consideration.

For any expected distribution \( \vec{a} \), the expected distribution of d-partitions can now be computed by multiplication with the matrix \( M \) and the number of languages/paradigms in the sample \( \ell \):

\[
\ell M \times \vec{a} = (\sum_{i=1}^n M_{\ell_1,m_i} a_m, \sum_{i=1}^n M_{\ell_2,m_i} a_m, \ldots, \sum_{i=1}^n M_{\ell_{B_c},m_i} a_m)
\] (5)

The equal distribution is a specific application of result (5), namely the case of \( a_i = 1/n \) for all \( i \in \{1, \ldots, n\} \).

For illustration, consider again Model 2: Then the matrix \( M^h \) has two rows and the following values where the second row contains only 0 except for columns 13 and 15:

\[
\begin{pmatrix}
(1-h)(1-2h)(1-3h) & h(1-h)(1-2h) & \cdots & h^2(1-h) & \cdots & h^2(1-h) & h^2(1-h) & h^3 \\
0 & 0 & \cdots & 0 & \cdots & 1-s & 0 & s
\end{pmatrix}
\]

Finding the best fitting \( h \) for the equi-parametrization now requires minimizing the following term for \( M^h \) as just given above (recall that \( \ell \) is the number of paradigms sampled):

\[
\text{fit}(h) = \left| \ell M^h \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix} - O \right|^2
\] (6)

One main effect of adding \( m_2 \) to the parameter space is that because of the equi-parametrization only 50% of the samples are assumed to have m-partition 1, rather than 100%. For low values of \( h \), we expect model 2 therefore to be better than model 1. For high values \( h \), however, the expected frequency of paradigm 1 is going to be less than the observed, and here model 1 will have an advantage. These intuitions are confirmed by our results: The dashed line in figure 5 shows how \( h \) affects the fit of model 2. Using again the optimize-function of R, we determine that the best fitting accidental homophony rate \( h \) for model 2 is 24.39%. The best fit made possible by model 2 is better than that of model 1. However, we will see below that model 2 still needs to be rejected.

### 3.5. Best Fit Parametrization and Model 2′

The second possibility we consider is the best fit distribution where we determine the assumed distribution from the actual data so as to minimize the difference between the expected and
observed distribution. Though the equal distribution is preferred and our investigation in the following is going to mostly use it, the best fit distribution is a useful comparison. Consider first the problems of the best fit parametrization, and then the ways it can be useful. The best fit distribution raises the problem of overfitting the data: the resulting morphological models have two parameters: $h$ and the distribution of parameterizations, which is itself a vector whose length $k$ is equal to the number of possible m-partitions and has $k - 1$ degrees of freedom. Since our investigation rests on only one data point – the actual distribution of languages – and there cannot be more than this one data point, the number of variables of the best fit model is likely to make the models too permissive. However, the best fit model is useful when it rejects a morphological system: Clearly if even the overly permissive best fit parametrization doesn’t allow us to fit the actual e-distribution well, a morphological parameter space can safely be rejected.

The mathematical computation involved in the best fit parametrization is more complex than the equal distribution models. Furthermore the result of the optimization depends on the accidental homophony rate $h$. We consider $h$ fixed in the following, though, because we can actually find the best fit parametrization for each $h$ within the minimization of $h$. The optimization problem is then specified by the term in (7). Because the likelihoods of all paradigms possible must add up to 1, one component of the distribution vector $\vec{p}$ is determined by the other $n - 1$ components. Here we assume that the last component is the dependent one.

$$\text{(7)} \quad \text{fit}(p_1, \ldots, p_{n-1}) = \ell M(h) \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_{n-1} \\ \ell - \sum_{i=1}^{n-1} p_i \end{pmatrix}^2 - O$$

There is an additional constraint on $\vec{p}$ that needs to be incorporated. Namely, none of its components may be negative since they represent probabilities. But it is frequently the case that the minimum of (7) has negative components: If the d-partitions in position $i$ occurs not at all or very rarely but is available as m-partitions, a negative value of $p_i$ in a sense would counterbalance the positive contribution to the $i$-th d-partition made by a frequent m-partition via accidental homophony. To rule out negative solutions, the values $p_1, \ldots, p_{n-1}$ must each be greater or equal to 0, and furthermore $\ell \geq \sum_{i=1}^{n-1} p_i \geq 0$ must be greater or equal to 0.

The following condition restates these requirements as an equation in the form $Ap + b \geq 0$ as required by the function $\text{constrOptim}$ of the software R.

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ -1 & -1 & -1 & \cdots & -1 \end{pmatrix} \vec{p} + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ -\ell \end{pmatrix} \succeq \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \text{(8)}$$

Finally, we need to determine the gradient of the function $\text{fit}$. The gradient is useful to find the minimum of a function quickly with numerical algorithms. In general, the gradient $\nabla f$ of a function $f$ taking $k$-long vectors as arguments is a function returning also $k$-long vectors that describe the slope of the surface $f$ defined by $f$ in a $k + 1$ dimensional space. Intuitively speaking, the gradient is useful to find the minimum because local minima have slope 0 and in other points the slope gives us information in which direction a local minimum can be found. In the present case, the function $\text{fit}$ may have many components and therefore it would be
computationally very inefficient to search minima without knowing the gradient. To compute
the gradient, note that $\text{fit}$ is the concatenation of the quadratic function $f(x) = x^2$, the linear
function $g(x) = Mx - O$ and the function $h(p_1, \ldots, p_{n-1}) = (p_1, \ldots, p_{n-1}, 1 - \sum_{i=1}^{n-1} p_i)$. Since the
derivative of $(f \circ g)'(x) = f'(g(x))g'(x)$, the gradient can be easily determined as the following,
where the $n \times (n - 1)$ matrix with all 1 on the diagonal, all -1 on the last row and 0 everywhere
else is the derivative of $h$:

$$\nabla \text{fit}(p_1, \ldots, p_{n-1}) = 2(M^T (-O)M)$$

As an example of a best fit parametrization, consider model 2' introduced above. For
model 1 evidently the best-fit parametrization would not differ from the equi-parametriza-
tion since there is only one parameter choice. For model 2', however, we expect the best fit
parametrization to balance the effects of the two possible m-partitions. For the discussion,
recall that $a_1$ be the parametrization for m-partition 1. Looking only at d-partitions 1 and 13, m-
partition 1 requires $h$ to be such that $a_1(1-h)(1-2h)(1-3h)$ is close to 18%, the percentage
of the observed data with d-partition 1. From m-partition 2 on the other hand, $(1 - a_1)(1-h)$
should also be close to 18%. Intuitively model 2' should improve on model 2 if $h$ is reduced
from 27%, and $a_1$ is increased above 50% because the the contribution of m-partition 1 to d-
partition 1 depends negatively on $p^3$. For example, with $h = 20\%$, the two expressions $a_1(1-
h)(1-2h)(1-3h) = 0.192a_1$ and $(1-a_1)(1-h) = 0.8 - 0.8a_1$ are equal for $p = 0.806$. So if
$h = 20\%$, the two approximation requirements mentioned above would be best when about 80%
of the paradigms are assigned to m-partition 1. When we take all 15 d-partitions into account in
a numerical calculation, the best-fit values turn out to be $h = 19.5\%$ and $p = 76.0\%$ for model 2'.
The dot line in figure 5 shows how for different $h$ values, the fit of the best-fit parametrization of
model 2' compares to the fit of model 1 and model 2 with the equi-distributed parametrization.
For logical reasons, the fit of Model 2' cannot be worse than Model 1 or Model 2 since both of
them are special parametrizations of 2'. Actually, the best-fit parametrization generates a better
fit than either Model 1 or 2. However, at this point we can’t yet say what would constitute a
satisfactory fit. In the next section, we address this issue.

3.6. The Likelihood of the Observed

In this section, we develop the method to evaluate the approximations introduced in the pre-
vious sections in theory independent terms. The starting point is to recall that the result of the
computations introduced above is in each case a probability distribution. We have presented
methods to adjust the syncretism rate $h$ so that the expected distribution of languages is as
close as possible to the observed distribution. In this section, we discuss how to determine
what likelihood a model assigns to an event like the observed distribution. Before we introduce
that in detail, we discuss how to compare models based on such a likelihood assigned to the
observation.

The line of statistical reasoning we employ is called Maximum Likelihood Modeling.
The perspective is substantially different from the statistics of hypothesis testing that is used in
psycholinguistic studies. In psycholinguistics, the focus is usually on showing that a null hypothesis can be ruled out. Typically, a psycholinguist aims at a demonstration that one variable that varies across the set of observations considered predicts the experimental data better than is predicted by the null hypothesis—most often the null hypothesis is simply that the variable shouldn’t affect the observations at all. To this end, a psycholinguist computes, for the model with the null hypothesis, how likely an outcome such as the observed outcome was, and rules against the null hypothesis if the computed $p$-value is low. If the null hypothesis is rejected, this then usually argues for an alternative hypothesis where the variable should affect the observation. Since the ideal experimental design used is one that is precisely tailored to only distinguish the null hypothesis from the alternative hypothesis, usually the alternative hypothesis is corroborated whenever the null hypothesis can be ruled out.

This type of reasoning also plays a role in our investigation, but a more limited role than is typically the case. Specifically, we want to rule out the two quasi null hypotheses we introduced above: the All-Accidental and the All-Parameters model. However, the data from syncretism pattern frequencies are much richer than the data gathered in a typical psycholinguistic experiment. The real challenge turns out to be to find a principle based hypothesis that is not ruled out by the observed data. But to accept an alternative hypothesis, the $p$-value for it should be high. So, while low $p$-values for the null hypotheses support a model that diverges from it, that model itself then also needs to have a high $p$-value since otherwise we would also feel compelled to rule out the alternative hypothesis. A situation, that the alternative hypothesis as well as the null hypothesis is ruled out by the observations can also occur in psycholinguistic experiment. For example, a predicted difference may go ‘in the wrong direction’ as follows: assume the alternative hypothesis predict that one set of observations should be greater than another set and the null hypothesis predict them to be not different. Then if we actually observe that the first set is on average smaller than the second set of observations, this is not consistent with either hypothesis. In actual psycholinguistic work, however, such an outcome seems to occur rarely as far as we know.

Both types of reasoning require the computation of the probability of a set of observation given a model. In our case, the observation is the observed distribution of patterns in Cysouw’s sample. Of course, though the probability of any single datapoint is always vanishingly small. The interesting probability is not only the one of this single point, but the $p$-value: that of all points that are at least as far away from the expected distribution as the datapoint under consideration – in our case, the datapoint from Cysouw’s sample. The reasoning is transparent in the example of the dice count we mentioned already in the introduction: imagine you don’t know how many dice I rolled, but you know that I rolled a 139. Does this allow you to reject the assumption that I rolled 40 dice (i.e. the 40-dice model)? The likelihood of rolling a 139 with 40 dice is about .036, i.e. below the .05 level. But clearly that should not lead to rejecting the 40-dice model: actually 140 is the most likely expected result when rolling 40 dice, and 139 is very close to that. The likelihood of being 1 or more away from 140 is greater than .96, and for this reason we certainly cannot reject the 40-dice model. On the other hand, the assumption that I rolled 139 dice all showing a one, the 139-dice model, can be rejected, because only the event of rolling a $6 \times 139 = 834$ is at least as extreme as 139 on the 139-dice model, so the $p$-value of 139, i.e. the likelihood of an outcome as extreme as 139, is smaller than $10^{-100}$ (namely, $2 \times 1/6^{139}$) on the 139-dice model.

What is an acceptable $p$-value for a model in maximum likelihood modeling? For hypothesis testing, the .05 level is established in the psycholinguistic community as the threshold that a $p$-value should be below in a published paper, though in some areas of neuroscience much lower $p$-values like $p < .001$ are standard. As far as we know, no corresponding threshold is
established for maximum likelihood modeling. The .05 level is relevant: If a model led to a $p$-value greater than .05, we would not be licensed to reject that model as inconsistent with our data. So this would be one outcome worth reporting. But there is another outcome worth reporting, which involves a relative improvement in $p$-value from a null hypothesis. For example, while a null hypothesis may predict the observed data to be close to impossible (e.g. a $p$-value of $10^{-20}$), an alternative hypothesis may predict a much higher likelihood for the observed data (e.g. $10^{-3}$). In this case, the enormous relative improvement in accounting for the observed data should be recognized as an important step towards an explanation of the observed data, though the .05-level is not yet reached. Therefore relative improvement of the $p$-values is as important as any absolute level. In the case of syncretism patterns specifically, it would be surprising if one alternative hypothesis about parametrization alone would completely explain the data since as we mentioned above the independence assumptions may be to some extent idealizations and the data may be difficult to model because of that.

In the remainder of the section, we now address the appropriate way to compute a $p$-value in our scenario. Unless $h$ is 0, the result of each model is a probability distribution where each d-partition is predicted to occur with some likelihood greater than 0. The observed outcome for the four cells of first person was shown above, and in other cases it will be similar: a distribution where some d-partitions occur a sizable number of times, possibly even greater the cap of 15 we applied above, while others occur very rarely or not at all. Two widely used tests for computing how likely an observed distribution across a number of cells is given a probability distribution: the Chi-Square test and the G-test.\textsuperscript{10} These tests provide a $p$-value for an observed distribution assuming a random event that can have finitely many discrete outcomes (or bins). In our case, the bins are the d-partitions. However, both cannot be used when the observed distribution contains 0s or even just counts below 5 for some bins.\textsuperscript{11} As we saw above that even with the full data available to us from 265 paradigms, only 5 of the 15 bins of in the first person case have count of 5 or greater. One might want to address this by gathering more data, however, this would require a massive effort in the case of eight cells, which holds the greatest interest: With eight cells, the number of bins is $B_8 = 4140$. In principle this might be possible, there probably are many more than the roughly 6000 languages recognized by Grimes (2000) once the independent status of so-called dialects is recognized, and in addition many languages have more than one relevant paradigm. Still, to gather a sample where all of over 4000 cells occurred at least five times is far beyond our current capabilities.

Since there is, as far as we presently know, no other statistical test applicable in our scenario, we are left with what is called in statistics the Monte Carlo method. This means we generate a random event that is distributed as our model predicts and count in a sample output how often an event at least as extreme as the real observation occurs. In the statistics software package R, the \texttt{chisq.test} defaults to such a simulation when given an input that doesn’t satisfy the requirement that all cells contain a count of at least 5 and we use this function in the end to compute the $p$-values for our data. In fact this choice already affected our choice of optimization procedure above: we minimized the term $\sum (E - O)^2 / E$ because this is the term the chi-square test depends on. For future research it is important to keep in mind that the distance metric used above needs to be chosen in accordance to the statistic test used at this point.

The Monte-Carlo method is slow method to estimate a likelihood, and this in particular

\textsuperscript{10} In the future, we plan to also apply the root-mean-square test of Perkins et al. (2011).

\textsuperscript{11} One way to address this problem would be to merge some of the bins so that all observed frequencies are at least 5. However, we have not explored this yet. For one, the results are strongly affected by the rebinning chosen. Furthermore, in a case with many 0s much of the power of the test is lost.
affects very low likelihoods. The Monte-Carlo method computes a large number of random events and counts how many of these are as far away from the expected distribution as the observed one. For a high probability this provides a useful estimate: for example, if 3500 of one million trials turn out to count, then 0.0035 is highly likely to be a good estimate of the $p$-value. For very low likelihoods, it is interesting to know whether they are $10^{-10}$ or $10^{-20}$. But with only one million trials, we are not able to distinguish such low likelihoods because for neither one are likely to encounter even one event that counts. All we could conclude from one million trials in this case, would be that the likelihood is below $10^{-6}$, i.e. the likelihood of one in a million. At this point, our computational resources don’t allow us to be more accurate than this. All we can say is that all the three models considered so far assign a very low likelihood to the observed distribution of languages $O$. The computation of two million trials underlying each line of the following table requires about 10 minutes per model on an Apple iMac personal computer with a 3.06 GHz Intel Core 2 Duo microprocessor.

<table>
<thead>
<tr>
<th>model</th>
<th>$h$</th>
<th>fit($h$)</th>
<th>$p$-value of $O$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>28.55%</td>
<td>221.5</td>
<td>$&lt; 10^{-6}$</td>
</tr>
<tr>
<td>Model 2</td>
<td>24.39%</td>
<td>179.5</td>
<td>$&lt; 10^{-6}$</td>
</tr>
<tr>
<td>Model 2'</td>
<td>25.12%</td>
<td>122.0</td>
<td>$&lt; 10^{-6}$</td>
</tr>
</tbody>
</table>

The low $p$-values that all three models assign the observed distribution entail that all three models can be rejected. Looking at the fit($h$) in the table and also in figure 5, we see that model 2 is better than model 1. But we can’t compute the improvement of model 2 relative to model 1 in terms of $p$-value: even model 2' where the parametrization is made to best-fit the observed data still assigns to the observed distribution a likelihood below one in a million. The rejection of model 1 is particularly interesting since it is the instantiation of the all-accidental model in this case, i.e. the general model Cysouw (2003) advocates.

Several further models are interesting and at this point everything is in place to just test them. We list the relevant values for four more models in the following table, where the Monte Carlo testing was only done with 20 000 repetitions. Model $∀$ is the all-parameters models. This model, as we mentioned, can also be considered a null hypothesis that makes only very minimal assumptions about grammar. The fact that it just like Model 1 can be safely rejected, shows that universal grammar constrains the parameter space of m-partitions in a non-trivial way.

<table>
<thead>
<tr>
<th>model</th>
<th>$h$</th>
<th>fit($h$)</th>
<th>$p$-value of $O$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model $∀$</td>
<td>1.2%</td>
<td>122.7</td>
<td>$&lt; 10^{-4}$</td>
</tr>
<tr>
<td>Model $&gt; 14$</td>
<td>9.6%</td>
<td>18.3</td>
<td>$=.190$</td>
</tr>
<tr>
<td>Model $&gt; 2$</td>
<td>5.5%</td>
<td>15.3</td>
<td>$=.106$</td>
</tr>
<tr>
<td>Model $&gt; 1$</td>
<td>4.1%</td>
<td>24.0</td>
<td>$=.009$</td>
</tr>
</tbody>
</table>

Models $> 14$, $> 2$, and $> 1$ all compute the parameter space directly from the observed distribution in figure 4. Namely, the parameter space consists of the m-partitions that as d-partitions have an observed frequency above the threshold of 14, 2, and 1 respectively. We expect these models to assign a high likelihood to the actual distribution, and this is in fact the case. This shows that some analyses that combine a parameter space restricted by universal grammar and some accidental homophony are consistent with Cysouw’s dataset.
4. Conclusions and Outlook

The distribution of patterns of identity and non-identity among paradigm cells (our d-partitions) across the world’s languages is unquestionably not random. Yet accidental homophony introduces an element of noise into these distributions, making it hard to distinguish the accidental from the systematic. In this paper, we have proposed a new approach to this problem, using statistical techniques: Syncretism Distribution Modeling.

Our main thesis extends a familiar premise using more precise mathematics: We treat accidental homophony as a random event in the statistical sense. Thus the difference between the systematic and the random should emerges in the formal analysis of large datasets. In section 3, we showed how to compute the p-value of an observed distribution of (d-)partitions of the paradigm space, assuming a given morphological model M. The p-value as computed in the 3.6 provides us with a direct way to interpret the fit between model M and observation O: Assuming model M, an outcome at least as extreme as observation O has likelihood p of occurring. By iterated application of this procedure, we can compare different morphological models by their fit with an observed cross-linguistic distribution of d-partitions. In this situation, we can then apply maximum likelihood modeling to determine the model or models that assign the highest likelihood to the actual observation O.

For reasons of space and time, in this paper, we described only the initial steps in this project, offering in essence a proof of principle. We focussed on the case of person paradigms, and at that mostly on the sub-paradigm created by the four cells of the first person, using an existing dataset (Cysouw 2001). The main linguistic result we report here is that analyses that assume a parameter space unconstrained by linguistic considerations (the models 1 and ∀ above) cannot explain the observed frequencies of d-partitions in this case. In ongoing work, we supplement these methods with an automatic procedure to compute a parameter space from a set of features. This gives us a way to directly assess the goodness of fit of different feature sets. For example, we may successfully infer a particular model (i.e. universal feature inventory) provides a good fit for the observed data – in fact one that converges with semantically plausible models assumed in the theoretical morphological literature – using this methodology. In addition we can compare the effect of allowing explicit ordering of the features or only Paninian implicit order. Our ongoing work addresses a question that we can only hint at here: in what way the model should be constrained? Here we pointed out that some parameter spaces (Model > 14 and Model > 2 above) fit the observed distribution well, but this is clearly a post-hoc result. The level of parameter spaces is for two reasons not the right level to look for satisfactory morphological models. For one, there are over 16 000 potential parameter spaces for the case of 4 cell paradigms, and given this large number, it is less surprising that some of them fit the observed distribution. Secondly no theory of grammar that we know of assumes that the parameter space is specified as a set of m-partitions. Rather the set of m-partitions that characterize a parameter space are only a derivative concept and the underlying assumptions are about features and the possible ways features can be related to morphemes.

A further step to be taken is to scale syncretism distribution modeling up to examine the full eight cell paradigm. While the principles are identical, scaling up turns out to be computationally challenging and might require more powerful computers than we can presently access. We conclude by noting that other four cell selections from the full eight cell paradigm exhibit similarly marked distributions as the first person selection that we focussed on. The d-partition counts of three other selections are shown in figure 6. For each, the pattern in the top left corner of each diagram shows by shading which cells were selected. The leftmost diagram shows the counts for the four non-first person cells. The center concerns the minimal cells, and the right
shows a selection of cells that don’t have any natural semantic category corresponding to them. The diagrams show that the person distribution is unusual also outside the first person domain.

Appendix on a Statistical Point

For each model, both the optimal homophony rate \( h \) and the p-value depend on the distance measure used (see the discussion in 3.3 and 3.6). In the following table, we present the results from a separate computation using two common distance measures for all the models presented in the text (except for model 2'): the Pearson’s chi-square approximation \( \sum (E - O)^2 / E \), and the Euclidean distance \( \sqrt{\sum (E - O)^2} \).\(^\text{12}\) We completed these computations in about one day of computation time while the final version of the paper was in production. The Monte-Carlo procedure to establish the p-values used 1 million tests each.

<table>
<thead>
<tr>
<th>model</th>
<th>Pearson optimal ( h )</th>
<th>Pearson p-value of ( O )</th>
<th>Euclidean optimal ( h )</th>
<th>Euclidean p-value of ( O )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>28.55%</td>
<td>( 10^{-6} )</td>
<td>22.16%</td>
<td>( 2 \times 10^{-6} )</td>
</tr>
<tr>
<td>Model 2</td>
<td>24.39%</td>
<td>( 10^{-6} )</td>
<td>27.1%</td>
<td>( 10^{-6} )</td>
</tr>
<tr>
<td>Model ( \forall )</td>
<td>1.20%</td>
<td>( 10^{-6} )</td>
<td>6.47%</td>
<td>( 10^{-6} )</td>
</tr>
<tr>
<td>Model &gt; 14</td>
<td>9.63%</td>
<td>0.20</td>
<td>2.01%</td>
<td>0.862</td>
</tr>
<tr>
<td>Model &gt; 2</td>
<td>5.51%</td>
<td>0.26</td>
<td>3.15%</td>
<td>0.101</td>
</tr>
<tr>
<td>Model &gt; 1</td>
<td>4.13%</td>
<td>0.075</td>
<td>4.98%</td>
<td>0.005</td>
</tr>
</tbody>
</table>

In general, both distance measures give similar results. The greatest discrepancies occur with model \( \forall \) and model > 14. This is expected because the Euclidean distance measure is informally speaking more sensitive to approximating the frequency of the highest frequency paradigms than the Pearson distance measure. Hence the Euclidean distance prefers for model > 14 a smaller rate of accidental homophony than the Pearson distance. As expected, the pattern is the opposite with model \( \forall \): the Euclidean prefers a greater accidental homophony rate than the Pearson distance.

\(^{12}\) We are still in the process of applying a Kolmogorov-Smirnov distance measure as a third measure. Note that the log-likelihood based measure are not applicable in our case since \( \log (O / E) \) doesn’t have a result whenever \( O = 0 \) or \( E = 0 \), and the former occurs in our data.
References


1. Introduction

Recent work exploring the evolution of human language (Berwick and Chomsky 2011; Chomsky 2005, 2007, 2010) presents the view that it is a one-time event that yielded unbounded Merge, with externalization of language as a subsequent secondary process that does not involve a genomic change. This hypothesis is embedded in the following version of the strong minimalist thesis (SMT) formulated by Chomsky (2010), slightly modifying the title of Sauerland and Gärtner (2007):

(1) Interfaces + Merge = Language

where language is understood to mean the faculty of language in the narrow sense (FLN) in the wording of Hauser, Chomsky, and Fitch (2002). Left out of this scenario is the status of Agree, which defines one of the two syntactic relations in a rather parsimonious computational system posited in the Minimalist Program (Chomsky 2004, 2008), the other one being set-membership given by Merge. What I would like to do in this article is to clarify how Agree fits the Merge-centered picture of language evolution. An obvious step is to link it to the process of externalization, but the problem is that externalization is regarded as a secondary process under the hypothesis summarized above. Is the emergence of Agree also a secondary process? My proposal is that Merge itself lays the foundation for the major players in the operation of Agree, namely, uninterpretable features, opening the door for externalization at the same time. In relation to this, I will also put forward a new conception of the Agree operation and pursue some of its consequences.

As an initial step in pursuing the idea, I begin in section 2 by summarizing the key properties that the Transfer operation must deal with in handing over the output of narrow syntax to the system that leads to the sensorimotor (SM) interface, called externalization above, focusing on uninterpretable features that drive the operation of Agree. A new conception of the Agree operation is presented next. The inventory of uninterpretable features and their theoretical status are also discussed. Section 3 applies the new conception of Agree to a rather rich empirical domain involving complex nominal morphology found in Jemez and Kiowa, reworking Harbour’s (2011) account. The result is a unified treatment of inverse in the two languages, which Harbour’s original account fails to achieve.

2. Agree and Externalization of LOT

FLN is a computational system that connects the conceptual-intentional (CI) interface to the SM interface, with narrow syntax sitting in its center. Lexical items are atomic units of narrow syntax computation, combined by the operation of Merge to form complex structure.
Semantic and phonological features are packed together in lexical items, to be used in the CI and SM systems, respectively. FLN must therefore include an operation of Transfer, which teases apart semantic and phonological information, sending the former to the CI interface and the latter to the SM interface.

It is not just phonological features that are handed over to the so-called PF branch of FLN, however. Narrow syntax contains an operation of Agree, which gives a value to originally value-less uninterpretable features by matching them with valued interpretable counterparts (Chomsky 2001). Once valued, uninterpretable features may be realized phonologically, a rather interesting property of natural language. This means that Transfer also picks out uninterpretable features that have undergone Agree and sends them to the morphological computational system that assigns to them phonological forms as in the theory of Distributed Morphology (Halle and Marantz 1993 and subsequent work). Thus, there are two sources of phonological information, namely, phonological features of lexical items and uninterpretable features valued by Agree.

What would be the form of FLN if we adopt the view put forth by Berwick and Chomsky (2011) and Chomsky (2005, 2007, 2010) that the system of externalization came later in the evolution of human language? They claim that FLN functioned as a system for a language of thought (LOT) at the earliest stage when human language emerged. This consequence is innocuous as far as lexical items are concerned. The research in the 1980’s emphasized the significance of empty categories, which are not realized phonetically, as strong evidence for the working of universal principles. Functional categories also often come without phonological content, perhaps reflecting the LOT nature of FLN. So it is perfectly natural to have a structure building system that dispenses with the link to the SM interface.

The computational system without externalization, however, cannot retain the operation of Agree in the form currently assumed. Agree outputs valued uninterpretable features that need to be taken care of by morphology. Without externalization, such features cannot exist. One might imagine handling them by brute-force deletion, but then, there is no point in throwing them into the computational system in the first place. Only with the potential of phonetic realization do uninterpretable features serve the purpose of establishing a link between the CI and SM interfaces through Agree, which matches interpretable features (CI entities) with their uninterpretable counterparts (see Watanabe 2008).

There are only two logical possibilities: the initial computational system of FLN is equipped with the Agree operation together with uninterpretable features, or it isn’t. I lean toward the affirmative answer here. At the same time, evidence for the primacy of the CI interface is considered to be robust, including the design optimization keyed to the CI interface and the lack of substantial differences between spoken and sign languages. How can one reconcile the CI primacy of FLN with the eventual linkage of Agree to the SM interface? I am not prepared to give a full story in this article, but hope to provide a plausible outline. I will concentrate on elucidating the nature of uninterpretable features by developing the suggestion made in Watanabe (2009) that they are recruited from semantic features. This suggestion locates the origin of uninterpretable features in the CI systems. I would like to elaborate it in this section.

2.1. Formal Features, Semantic Interpretation, and Agree

Uninterpretable features that drive the operation of Agree are classified as formal features in the typology of features proposed by Chomsky (1995), according to which there are two kinds of features that contribute to semantic interpretation in the CI systems. One is called semantic features, and the other interpretable formal features. The difference between the two is that semantic features are inert during narrow syntax whereas formal features enter into agreement when interpretable ones are matched with their uninterpretable counterparts.
under appropriate locality conditions. The proposal of Chomsky (2001), mentioned above, views this process as valuation.

Crucially, there are no formal features like [ ± consonantal]. For this reason, Watanabe (2009) argues that formal features are recruited from resources on the CI side even though there are uninterpretable versions of them. Furthermore, comparative study of different stages of English suggests that there is a parameter concerning feature classification allowing the definiteness feature to act as formal in some languages and as semantic in others, reinforcing the case for the idea that formal and semantic features share a common evolutionary origin. Then, precisely how are uninterpretable formal features created out of CI resources?

To answer this question, we need to sharpen the notion of uninterpretable features that receive a value through the Agree operation. For this purpose, we can make use of Harbour’s (2007, 2011) proposal that uninterpretable features are characterized by the presence of both the plus and minus values, as in [+F, –F] for the feature F. This specification is uninterpretable because it gives contradictory instructions. Now, observe that it is a small step from this conception of uninterpretability to identification of the evolutionary process that derives it: binary Merge. Harbour himself (2011: note 5) comments that given a combinatorial system such as syntax, the existence of [+F] and [–F] can automatically give rise to [–F, +F] unless a further stipulation blocks cooccurrence of the two. Since Merge forms a set, the uninterpretable version of F is \{+F, –F\}, strictly speaking.

The prerequisite for this hypothesis is that conceptual resources for formal features must come in pairs. Harbour (2011) indeed argues for the bivalent nature of number features on the basis of the analysis of Kiowa and Jemez, which will be revisited in section 3. See also Watanabe (2010) for the same conclusion from different empirical material and Watanabe (to appear) for the binarity of the person feature system. Quite generally, the strictly privative feature system is incompatible with the use of valuation for agreement, as Adger and Svenonius (2011) point out, since valuation presupposes two distinct states of a feature. After the privative system is ruled out, there remain various logical possibilities for multi-valued features. Only the binary system, however, is compatible with the idea that Merge creates uninterpretable formal features.

Once uninterpretable features are made available for the FLN computation and are actually thrown into the derivation together with lexical items, the need to get rid of them immediately arises. The outlet must be the SM systems. Thus, we can conclude that the motivation for externalization comes from exploitation of the CI resources by the operation of Merge. One may say that this is another respect in which the CI interface enjoys the primary status. FLN is destined to develop the mechanism of externalization even though it is an instrument for LOT. This reasoning also suggests that the distinction between semantic features and interpretable formal features is superfluous and derivative. Uninterpretable features are created out of interpretable ones and therefore can always be matched with them (with one important exception that will be taken up in section 2.3).

It is an open empirical question whether the earliest stage of FLN already possessed the system for externalization. The availability of uninterpretable features could be only latent at the beginning. The question needs to be addressed together with whether the earliest stage of FLN already possessed the operation of Agree. One relevant consideration has to do with the fact that Agree serves as a search mechanism for the proper operation of Internal Merge. It is not clear whether Internal Merge can work in an optimal way without the help of Agree. Did

---

1 I am not saying that Merge forms uninterpretable features as part of derivation. I assume that they are given for each individual language, with the precise inventory parametrized. My claim is that binary set formation is responsible for the phylogenetic origin of uninterpretable features.

Let me add that recursion is not needed for the creation of uninterpretable features. It would only give rise to monsters that the externalization system cannot handle.
the earliest stage of FLN lack Internal Merge, even though this mode of Merge is given for free (Chomsky 2004)? I will not resolve the issue here. Instead, I now turn to the formulation of Agree.

2.2. A New Theory of Agree

The idea that uninterpretable features come in the form of \{+F, –F\} allows us to reconsider the valuation mechanism of the Agree operation. Harbour (2007, 2011) takes the view that values matched with those of the interpretable counterpart are retained and become visible at PF, whereas unmatched values get deleted (or simply become invisible at PF). Under this conception, retention of the matched values is the core process of valuation.

I would like to propose instead that valuation is indeed deletion. The traditional intuition tells us that what happens in agreement is something like (2).

(2) a. Valuation of \{+F, –F\} with +F means deletion of –F, producing \{+F\}.
   b. Valuation of \{+F, –F\} with –F means deletion of +F, producing \{–F\}.

But if valuation deletes one of the two values, there is no logical necessity that things should work as in (2). Agreement can operate in the opposite direction, as in (3).

(3) a. Valuation of \{+F, –F\} with +F means deletion of +F, producing \{–F\}.
   b. Valuation of \{+F, –F\} with –F means deletion of –F, producing \{+F\}.

In other words, matching of \{+F, –F\} with +F (or with –F) itself does not dictate which value should go as a result of valuation. No matter which value is deleted, it is a stipulation made arbitrarily. And as long as valuation always involves deletion, this arbitrary choice must be made for each case of agreement, a locus of parametrization.

Is there any evidence that the option in (3) is needed? Actually, a mundane fact about English subject-verb agreement seems to require it. Commenting on data like (4), Chomsky (1975:234) remarks that the singular form of the verb is matched with the plural subject and vice versa, viewing the suffix -s as a marker of [–singular] both in (4a) and in (4b).

(4) a. The boys like apples.
   b. The boy likes apples.

The conventional treatment leaves the two -s suffixes in (4a, b) as completely unrelated. The new conception of agreement allows us to accommodate data patterns like (4) in a uniform account of what looks like a single morphological piece.

Option (2) is also needed, in view of the following Spanish data from Harris (1991:34):

(5) a. Mi sobrino es alto.
    my nephew is tall.m
   b. Mi sobrina es alta.
    my niece is tall.f

Here, the feminine subject forces the feminine form of the adjective, and the masculine subject the masculine form.\(^2\) Given the conventional treatment of agreement, one may be led to think that option (2) is the unmarked choice. But the matter is entirely empirical. Future research should examine the distribution of options (2) and (3).

Anders Holmberg pointed out to me that the mechanism of valuation in terms of deletion

\(^2\) See Watanabe (2011) for an account of agreeing predicative adjectives in terms of Multiple Agree.
finds an analogue in pairing of yes-no questions and answers. Yes-no questions denote a set of propositions of the form \{p, \neg p\}, from which the answer is formed through deleting one of the propositions. Note the striking formal similarity. Valuation may have been designed on the model of the CI system that lies behind yes-no questions and answers.

2.3. CI Resources

Let us examine next whether the conceptual resources for the \(\phi\)-feature system were available at the birth of FLN.

To start with the number features, Hauser, Barner, and O’Donnel (2007) report that rhesus monkeys exhibit the sensitivity to the singular-plural distinction. There does not seem to be any obstacle to the assumption that \([\pm\text{singular}]\) was recruited from pre-linguistic conceptual resources. For \([\pm\text{augmented}]\), on the other hand, there is no comparative study, as far as I am aware. Watanabe (2010) suggests that \([\pm\text{augmented}]\) is involved in licensing of numerals, linked to the natural number system. If the notion of natural numbers was made possible by the emergence of Merge along the lines suggested by Chomsky (2008), \([\pm\text{augmented}]\) could not have existed prior to the appearance of FLN, at least in the form linked to natural numbers. The notion behind \([\pm\text{augmented}]\), though, is simply the contrast between minimal and non-minimal, informally speaking. It is conceivable without assuming language. The rationale connecting \([\pm\text{augmented}]\) to natural numbers in Watanabe’s (2010) discussion is that the hierarchical specification in the shape of \([-\text{augmented}]\) \([+\text{augmented}]\) serves as the successor function by yielding the minimal element of the non-minimal part. It is possible that only the linkage to natural numbers (and hence numerals) was established by the emergence of FLN. We also need to address whether the contrast between minimal and non-minimal is a notion available for non-human animals. The singular-plural distinction, which rhesus monkeys are sensitive to, may be due to \([\pm\text{augmented}]\).

Turning now to the person features, a naïve reaction would be to say that notions like speaker and hearer are impossible without language. But person distinctions are deictic in nature. A member of a species facing another member is a commonplace event, especially for animals living in groups. The notions of speaker and hearer were probably transposed from situations of such social interactions. One may even talk about non-verbal communication. Only the verbal part is an innovation, then.

The interpretability of gender features itself is debatable in the first place, but Percus (2011) suggests that the problem can be sidestepped by providing an explicit interpretation procedure that specifically mentions the irrelevance of the masculine/feminine contrast for inanimate entities. He takes up Italian as the simplest case, which only has two classes of nouns. Though there does not seem to be any consensus on what are the possible noun classification systems for natural language, one can adopt a strategy similar to Percus’s for any classification system as long as it employs semantic notions as its basis. In principle, then, features of noun classification can be recruited from the CI systems. The real task is to identify and enumerate semantic notions that define possible noun classification systems. Chomsky (2007, 2010) stresses that concepts realized as lexical items are radically different in nature from those entering into representational systems of other animals. If any one of these concepts turns out to be crucially involved, that must be treated as an innovation. Much work remains to be done.

To sum up so far, at least some of the \(\phi\)-features can be attributed to pre-linguistic conceptual resources. One can therefore construct a primitive agreement system on that basis.

The account of \(\phi\)-feature agreement cannot be complete without taking up the status of case, however. And here, we face a real difficulty. Structural case has been considered to be a

---

3 For detailed discussion of various examples, see Chomsky (2000b).
quintessential example of uninterpretable entities in FLN. Even if one can point to some abstract semantic correlates of structural case (see Svenonius 2007), it is simply impossible to organize relevant concepts into binary features. In a way, this status of structural case is reflected in the valuation system of Chomsky (2001), where case is simply assigned a value as a by-product of φ-feature agreement, depending on the type of the probe, without being matched with an interpretable counterpart. Binarity requirements on conceptual resources put rather severe constraints on possible avenues for exploration. In this connection, inherent case, which is linked to thematic role interpretation (Chomsky 1986), must also be taken into account, because it is on a par with structural case in being reflected in morphological realization, at least in some languages. Important hints might be sought in the way externalization works. As shown by Hiraiwa (2010), realization of case is regulated by the OCP in a cyclic fashion. Given the theoretical origin of the OCP (Leben 1973), it might be fruitful to compare case with tone, though only future research can tell whether this promissory note can be cashed.

3. Inverse in Jemez and Kiowa

In the rest of this article, I would like to show how the new theory of agreement treats inverse marking in Jemez and Kiowa. One of the great achievements in Harbour’s (2011) work is to identify nearly identical sets of noun classes in Jemez and Kiowa. Previous work on Jemez such as Noyer (1997) classifies Jemez nouns somewhat differently, failing to recognize the unity of noun classification in Jemez and Kiowa. Harbour accomplishes a nearly complete unification by pointing to the special status of dual in Jemez. My account will inherit this insight. On the other hand, inverse is still characterized separately for Jemez and Kiowa under Harbour’s analysis, the difference between the two languages being treated as nothing but an arbitrary stipulation. I will show that the new theory of agreement allows us to capture the fact that inverse works in exactly the same way in Jemez and Kiowa, enabling us to make progress in this respect as well. The fact remains, though, that there is a difference between Jemez and Kiowa. The last key element in my proposal is the treatment of the special status of dual in Jemez pointed out by Harbour, which finds an independent cross-linguistic motivation, unlike the way Harbour handles it. Once the special status of dual in Jemez is correctly pinned down, the difference between the two languages follows.

The discussion proceeds in the following way. Section 3.1 reviews Harbour’s (2011) account, introducing at the same time the empirical material to be explained. Section 3.2 compares the new theory of agreement with Harbour’s execution of agreement. Section 3.3 presents a novel uniform analysis of inverse in Jemez and Kiowa, attributing the difference between Jemez and Kiowa to something other than inverse.


A very intriguing characteristic of Jemez and Kiowa is that the majority of nouns have an inherent number value and that when the actual number interpretation differs from the one expected from the inherent value, a special marker called inverse is suffixed to the noun. In this section, I will review how Harbour (2011) handles this phenomenon.

Jemez and Kiowa make a tripartite distinction of singular, dual, and plural in number. These number categories are defined in terms of elementary features as follows:

---

4 Let me add that regulation is intended to cover even those cases where multiple instances of identically case-marked phrases are tolerated by the OCP in a local domain, as in the Japanese nominative case.

5 See Maling, Yip, and Jackendoff (1987) in this connection.
Informally speaking, [+augmented] designates non-minimal entities. Since [±augmented] is interpreted relative to [±singular], the combination [–singular, –augmented] picks out minimal elements in the [–singular] entities, namely, dual. Similarly, [±singular, +augmented] is a contradictory specification for the purposes of semantic interpretation and therefore is impossible. For a model-theoretic definition of these features, see Harbour (2011).

The function of inverse marking can be observed in the following Kiowa examples:

(7) a. Tógúł Ø-dɔ́ɔ́.
   young.man 3sg-be
   ‘It’s a young man.’

b. Tógúł e-dɔ́ɔ́.
   young.man 3du-be
   ‘It’s two young men.’

c. Tógúúdó e-dɔ́ɔ́.
   young.man.inv 3inv-be
   ‘They’re young men.’

(8) a. Kútaaɗo e-dɔ́ɔ́.
   pencil.inv 3inv-be
   ‘It’s a pencil.’

b. Kútaa e-dɔ́ɔ́.
   pencil 3du-be
   ‘It’s two pencils.’

c. Kútaa gya-dɔ́ɔ́.
   pencil 3pl-be
   ‘They’re pencils.’

The suffixed noun in (7c) is interpreted as plural, whereas the bare form in (7a, b) means either singular or dual, depending on the verbal inflection. In (8), on the other hand, the suffixed form corresponds to singular, and the bare form either to dual or to plural. Note also that inverse-marked nouns trigger inverse agreement on the verb in (7c) and (8a). Harbour calls nouns like tógúł SDI nouns, and nouns like kútaa IDP nouns, to indicate the agreement type of each noun class for the singular-dual-plural triad in this order, where S is singular, D dual, P plural, and I inverse. The inherent number of SDI nouns is [–augmented], and that of IDP nouns [–singular]. Thus, noun classes are defined in terms of number features. Inverse marking is used when the intended interpretation refers to the opposite value: [+augmented] for SDI nouns and [±singular] for IDP nouns.

The Kiowa inverse can express dual, too, as illustrated in (9).

(9) a. Nɔ́ a-dɔ́ɔ́.
   I 1sg-be
   ‘It’s me.’

b. Nɔ́ e-dɔ́ɔ́.
   I 1inv-be
   ‘It’s me and him/her.’ or ‘It’s me and them.’

According to Harbour, the first person pronoun is the only item whose inherent number is
characterized as [+singular]. It exhibits the SII behavior. (9b), where the verb shows inverse agreement, is ambiguous between dual and plural for that reason.6 Kiowa also has nouns whose cardinality is expressed transparently by the verbal agreement marker.

(10) a. X!oú \(\theta\)-d\(\ddot{\text{o}}\).  
    stone  3sg-be  
    ‘It’s a stone.’

b. X!oú \(\epsilon\)-d\(\ddot{\text{o}}\).  
    stone  3du-be  
    ‘It’s two stones.’

c. X!oú \(\text{gya}\)-d\(\ddot{\text{o}}\).  
    stone  3pl-be  
    ‘They’re stones.’

This type of noun does not have an inherent number specification and lacks inverse marking. Nouns of this type belong to the SDP class.

Harbour (2011) presents (11) as the list of noun classes attested in Kiowa.

(11) Kiowa noun classes (Harbour 2011: 578)

<table>
<thead>
<tr>
<th>class</th>
<th>class features</th>
<th>semantic characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDP</td>
<td>Ø</td>
<td>default</td>
</tr>
<tr>
<td>SII</td>
<td>[+singular]</td>
<td>first person only</td>
</tr>
<tr>
<td>SDI</td>
<td>[–augmented]</td>
<td>independently mobile objects</td>
</tr>
<tr>
<td>IDP</td>
<td>[–singular]</td>
<td>vegetation; most non-SDI implements; most non-SDI body parts</td>
</tr>
<tr>
<td>IDI</td>
<td>[–singular, –augmented]</td>
<td>hair types; midsize fruit growing in clusters</td>
</tr>
<tr>
<td>SDS</td>
<td>[+group]</td>
<td>non-shape-inductive objects</td>
</tr>
<tr>
<td>IDS</td>
<td>[–singular, +group]</td>
<td>vegetation occurring in natural collections; implements that may act collectively</td>
</tr>
<tr>
<td>SSS</td>
<td>[+augmented, +group]</td>
<td>nongranular mass nouns</td>
</tr>
<tr>
<td>PPP</td>
<td>[+augmented, –group]</td>
<td>pluralia tantum nouns; granular mass nouns (for some speakers)</td>
</tr>
</tbody>
</table>

I will not bother the reader with further examples of each noun class. These are amply discussed in Harbour’s (2007) book. Our major concern in this article is how the inherent feature specification of each noun class is combined with the precise characterization of inverse to yield the attested agreement pattern. For this particular purpose, (11) suffices.

The list in (11) includes a feature [±group] that I have not explained yet. It plays a special role in agreement within DP in Harbour’s account. Its function will be taken up when Harbour’s featural characterization of inverse is reviewed below. Let me just say at this moment that the active role of [±group] is limited to pluralities (hence [+augmented]), where the presence ([–group]) or absence ([+group]) of salient subparts matters.

For Jemez, Harbour (2011) claims that its noun classification system can be summarized as in (12).

---

6 The first person pronoun is not suffixed with an inverse marker (Harbour 2007: 88). Kiowa has a null inverse suffix that is used for nouns ending in /p/ or /t/ (Harbour 2007: 55), but the first person pronoun does not fall under this generalization. This fact may be due to the different morphological status of the D head associated with pronouns. As will be explained below, the inverse marker is considered to be located at D.

Note also that Kiowa distinguishes between exclusive and inclusive. (9b) exemplifies exclusive dual and plural. Inclusive dual and plural also trigger inverse agreement (Harbour 2007: 82).
(12) Jemez noun classes (Harbour 2011: 580)

<table>
<thead>
<tr>
<th>class</th>
<th>class features</th>
<th>semantic characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIP</td>
<td>Ø</td>
<td>default</td>
</tr>
<tr>
<td>SII</td>
<td>[–augmented]</td>
<td>animates</td>
</tr>
<tr>
<td>IIP</td>
<td>[–singular]</td>
<td>vegetation; artifacts; body parts</td>
</tr>
<tr>
<td>III</td>
<td>[–singular, –augmented]</td>
<td>weakly granular mass nouns</td>
</tr>
<tr>
<td>SIS</td>
<td>[+group]</td>
<td>collective counterpart of SIP</td>
</tr>
<tr>
<td>IIS</td>
<td>[–singular, +group]</td>
<td>collective counterpart of IIP</td>
</tr>
<tr>
<td>SSS</td>
<td>[+augmented, +group]</td>
<td>nongranular mass nouns</td>
</tr>
<tr>
<td>PPP</td>
<td>[+augmented, –group]</td>
<td>granular mass nouns</td>
</tr>
</tbody>
</table>

The behavior of a noun from the IIP class is illustrated below.


     that-inv box-inv I 1sg-3inv-lie.sg/du
‘That box is mine.’
‘Those [two] boxes are mine.’

b. Nį́ t’ëtibæ nį́ i-Ø-g’ó-.

     that box I 1sg-Ø-lie.pl
‘Those [several] boxes are mine.’

The suffixed form in (13a) is ambiguous between the singular and dual readings here. Some more Jemez examples will be presented later on when I propose a new analysis.

It is quite remarkable that the set of noun classes in Jemez is featurally almost identical to the Kiowa counterpart. The only difference is the absence of the [+singular] class from (12). This near identity is not a coincidence. Both Jemez and Kiowa are members of the Kiowa-Tanoan family.

Another observation to be made is that when inverse marking is available for a class, dual is always marked with inverse in Jemez, unlike in Kiowa. In fact, nouns with the same inherent number specification display different agreement profiles in the two languages. The Jemez noun t’ëtibæ in (13), a member of the IIP class, is inherently marked with [–singular]. Kiowa nouns with the same inherent number specification, however, belong to the IDP class. Harbour (2011) attributes this difference to parametrization of the way inverse is characterized in terms of number features, a topic to which I will turn next.

Harbour (2011) posits the following structure for DPs:

(14)                D P
                      NumberP        D
                      Class (Noun)        Number

Class is the locus of inherent number specification, hosting a subset of {±singular, ±augmented, ±group}. These features are chosen according to the conceptual content of the noun in question and do not contribute to model-theoretic interpretation. The semantic characteristics associated with each class in (11) and (12), which do not concern us here, are regarded as relevant to inherent number specification. Number consists of [±singular, ±augmented], determining the number interpretation of DPs as in (6). Inverse marking is associated with the D head, which hosts [±singular, ±augmented] as uninterpretable features.

---

7 Membership of each class differs slightly, too. Yumitani (1998) also notes speaker variation for Jemez.
that need valuation. Crucially, Class and Number are assumed to jointly value D.

According to Harbour’s analysis, inverse marking arises under the following conditions:

(15) Inverse
   a. $\text{Inv} \iff [-F, +F]$ (Kiowa)
   b. $\text{Inv} \iff [\alpha F, \alpha G]$ (Jemez)\(^8\)

As discussed by Harbour (2007) and Yumitani (1998), the inverse marker within DP has allomorphs, including zero mentioned in note 6. Thus, Inv in (15a) and (15b) is nothing other than a cover term for various morphologically realized forms, as noted by Harbour (2011). (15a) and (15b) should be taken as abbreviations for actual Vocabulary items in the framework of Distributed Morphology (Halle and Marantz 1993, Embick 2010).

Let us now turn to the mechanism of valuation that can yield the feature specifications in (15). Chomsky (2001) hypothesizes that uninterpretable features lack values and therefore need valuation. Harbour (2007, 2011) proposes instead that uninterpretable features take the form of $[-F, +F]$, as we have already seen, and that they undergo valuation as described in (16).

(16) Valuation in agreement
   a. An uninterpretable feature is valued if and only if it is matched with an interpretable counterpart in the appropriate domain.
   b. A feature that has been matched is visible at PF. (Unmatched uninterpretable features are invisible.)

In other words, valuation functions as retention of the matched feature under this view.

To see how this mechanism works, consider the Jemez IIP noun $t\acute{e}t\ddot{h}b\acute{a}$ in (13), which is $[-\text{singular}]$. The input to agreement has the following options, where $uF$ is used as a shorthand notation for $[-F, +F]$ to save space:

(17) a. $[[\text{Class} -\text{singular}] [\text{Number} +\text{singular,} -\text{augmented}]] [\text{D} \ u\text{singular}, \ u\text{augmented}]$
   b. $[[\text{Class} -\text{singular}] [\text{Number} -\text{singular,} -\text{augmented}]] [\text{D} \ u\text{singular}, \ u\text{augmented}]$
   c. $[[\text{Class} -\text{singular}] [\text{Number} -\text{singular,} +\text{augmented}]] [\text{D} \ u\text{singular}, \ u\text{augmented}]$

The fact that both Class and Number value D means that the result of valuation is the union of the two sets of features. The resulting content of D is shown below.

(18) a. $[[\text{D} +\text{singular,} -\text{singular,} -\text{augmented}]] <= (17a)$
   b. $[[\text{D} -\text{singular,} -\text{augmented}]] <= (17b)$
   c. $[[\text{D} -\text{singular,} +\text{augmented}]] <= (17c)$

(15b) predicts that (18a) and (18b) lead to inverse marking while (18c) does not. This is the correct result, as can be seen from (13) or the class label of the noun, which is IIP. All the other Jemez nouns that do not involve $[\pm\text{-group}]$ are also accounted for correctly. Since dual is $[-\text{singular,} -\text{augmented}]$, the retention view of valuation makes sure that dual nouns will always be marked with inverse, given (15b), as long as there is no other interfering factor.

The presence of $[\pm\text{-group}]$ can constitute such an interfering factor. Harbour (2011: 572) posits special provisions in (19).

---

\(^8\) Since the features in question are $[\pm\text{singular}]$ and $[\pm\text{augmented}]$, a more accurate specification of the Jemez inverse under Harbour’s account should be $[\alpha\text{singular,} \alpha\text{augmented}]$. 

- 64 -
(19) Effect of $[±\text{group}]$ on valuation of D
   a. In presence of $[+\text{augmented}]$, $[+\text{group}]$ causes D to be valued as $[+\text{singular},
   -\text{augmented}]$ (i.e., as a singularity).
   b. In presence of $[+\text{augmented}]$, $[−\text{group}]$ causes D to be valued as $[−\text{singular},
   +\text{augmented}]$ (i.e., as a plurality).
   c. In absence of $[+\text{augmented}]$, $[±\text{group}]$ is inert.

Nouns inherently marked as $[+\text{augmented}, ±\text{group}]$ are either SSS or PPP, depending on the value of $[±\text{group}]$. For the classes of SIS and IIS in (12), the effect of $[+\text{group}]$ shows up only when Number contains $[+\text{augmented}]$, namely, in plural.

Consider Kiowa next. Combined with the mechanism of valuation in (16), (15a) predicts that inverse marking will be used whenever a value of the inherent number specification is in conflict with one of the features in Number, again, as long as $[±\text{group}]$ is absent. This prediction is borne out in (11). The effect of $[±\text{group}]$ appears in the same way as in Jemez.

This wraps up the review of Harbour’s (2011) account of Jemez and Kiowa. Despite the almost complete unification of noun classes, two different characterizations of inverse are offered for Jemez and Kiowa. This move is forced because the two languages differ in the distribution of inverse marking. To account for the difference, one must posit some source of variation, which under Harbour’s account is the characterization of inverse. Though the divergent treatments of inverse may not be necessarily problematic, I will show below that a single formulation of inverse suffices for the two languages. Two keys ideas are the new theory of agreement introduced in section 2.2 and an independently motivated way to explain away the difference between Jemez and Kiowa in the distribution of inverse.

3.2. Valuation in Agreement

In this section, I will elaborate on the new theory of agreement, especially, with regard to the mechanism of valuation. My proposal capitalizes on Harbour’s new conception of uninterpretable features, as mentioned above. Harbour presents his own theory of valuation on the basis of the assumptions in (16). In order to explore what other possibilities are available, let us reconsider how valuation works under Harbour’s analysis.

The kind of agreement involved in inverse marking is multiple Agree in the sense of Hiraiwa (2005), where a single probe is matched with multiple goals. Under the retention view on valuation adopted by Harbour (2011), (17a) leads to (18a), both repeated below.

(17) a. $[\text{Class }−\text{singular}] [\text{Number }+\text{singular, }−\text{augmented}]$ $[D \text{usingular, }u\text{augmented}]$

(18) a. $[D +\text{singular, }−\text{singular, }−\text{augmented}]$

Here, D is the probe, and Class and Number the goals. Note that deletion is involved in the process, if deletion is taken to be the means of rendering features invisible as described in the parenthesized portion of (16b). The algorithm for valuation is as follows:

(20) Valuation algorithm (retention version)
   a. Step 1: Form the union of the features contained in the goal(s).
   b. Step 2: Delete from the probe the feature(s) not contained in the union formed in Step 1.

---

9 And this is the only way of carrying out valuation, if we do not want to introduce an additional mechanism of stamping a validation mark on matched features that keeps them visible till Vocabulary Insertion takes place.
In the case of (17a), Step 1 gives \{+\text{singular}, –\text{singular}, –\text{augmented}\}. Deletion is carried out in relation to this set.

A conceptually simpler way of handling valuation is to do away with union formation. It is formulated in (21).

\[(21)\text{ Valuation algorithm (new version)}\]

Delete from the probe the feature(s) whose value is opposite to that of the features in the goal(s).

According to the algorithm in (21), one only needs to examine individual features of the goal(s) to carry out deletion. In the case of (17a), valuation therefore means deletion of \{+\text{singular}, –\text{singular}, and +\text{augmented}\}.

It should be observed that under Harbour’s account, the new conception of uninterpretable features plays no significant role in the valuation of the features at D. The same result can be obtained by simply copying feature values from the two goals at Class and Number, instead of following (16) or (20). The only role of the new conception in his account I can see is to remove from researchers’ mind psychological barriers to accepting \([-\text{F}, +\text{F}]\) as the characterization of inverse in Kiowa. The idea, I conjecture, is that since the probe starts out with \([-\text{F}, +\text{F}]\), there should not be anything wrong with \([-\text{F}, +\text{F}]\) being the result of valuation. The deletion algorithm in (21), on the other hand, captures effectively the essence of valuation under the new conception of uninterpretable features. If the probe is initially provided with \([-\text{F}, +\text{F}]\), valuation needs to take the form of deletion.

Let us now return to (17a) by adopting the official set notation for uninterpretable features, namely, \{-\text{F}, +\text{F}\}, whose conceptual foundation is given by Merge. The content of D prior to valuation is \{\{+\text{singular}, –\text{singular}\}, \{+\text{augmented}, –\text{augmented}\}\}. Deletion of \{+\text{singular}, –\text{singular}, and +\text{augmented}\} then results in \{\{\}, –\text{augmented}\}\}. \{\} is an empty set, which is an entity that has a special status in set theory, since it is a unique set that has no member.

More generally, the valuation algorithm in (21) produces the following results:

\[(22)\begin{align*}
a. \text{valuation of } \{-\text{F}, +\text{F}\} & \text{ with } +\text{F} = \text{deletion of } –\text{F}, \text{ producing } \{+\text{F}\} \\
b. \text{valuation of } \{-\text{F}, +\text{F}\} & \text{ with } –\text{F} = \text{deletion of } +\text{F}, \text{ producing } \{-\text{F}\} \\
c. \text{valuation of } \{-\text{F}, +\text{F}\} & \text{ with } +\text{F} \text{ and } –\text{F} = \text{deletion of } –\text{F} \text{ and } +\text{F}, \text{ producing } \{\}
\end{align*}\]

(22a, b) repeats (2a, b). Note that the empirical difference between (20) and (21) shows up only when multiple goals possess conflicting values as in (22c). That is why Jemez and Kiowa provide crucial evidence for this theory of agreement.\(^{10}\) With the new mechanism of agreement/valuation in hand, let us take a fresh look at the data.

3.3. A Uniform Treatment of the Kiowa-Tanoan Inverse

As the discussion in the previous section shows, the new theory of valuation produces different results for the uninterpretable features of D. This inevitably forces us to look for a new featural definition of inverse. Let us start pursuing consequences of the new theory by

\(^{10}\) The mode of valuation in (3) produces the same result as (22c) in the case of agreement with multiple goals having conflicting values. And (22c) is what matters in the account of Jemez and Kiowa. For the account of Jemez and Kiowa below, however, I will assume the mode in (2), since the alternative mode in (3) produces \{+\text{singular}, +\text{augmented}\} instead of (25b) for (24b). It is not obvious whether such a combination is allowed as the feature content of inflectional morphology, even though \{+\text{singular}, +\text{augmented}\} should in principle be ok as long as semantic interpretation is not at stake.
analyzing Kiowa, which does not involve an additional complication related to dual.

As promised above, I will stick with a single, unified characterization of inverse, which turns out to be (23).

(23) Inverse in Jemez and Kiowa

\[
\text{Inv} \Leftrightarrow \{ \} 
\]

This is the formal expression of the idea that inverse indicates the number information opposite to the inherent specification of the noun, under the theory of valuation formulated in (21). When the value of Class and that of Number differ with respect to [±F], valuation necessarily produces an empty set for that feature in D.

To show that (23) yields the correct result for Kiowa, let us consider an IDP noun illustrated in (8), repeated below.

(8) a. Kútaadɔ e-dɔɔ.  
    pencil.inv 3inv-be  
    ‘It’s a pencil.’

b. Kútaa e-dɔɔ.  
    pencil 3du-be  
    ‘It’s two pencils.’

c. Kútaa gya-dɔɔ.  
    pencil 3pl-be  
    ‘They’re pencils.’

Since IDP nouns in Kiowa are inherently marked with [–singular], we have the following representations for singular, dual, and plural:

(24) a. \( [[[\text{Class} –\text{singular}] [\text{Number} +\text{singular}, –\text{augmented}]] [D \text{usingular}, \text{uaugmented}]] \)

b. \( [[[\text{Class} –\text{singular}] [\text{Number} –\text{singular}, –\text{augmented}]] [D \text{usingular}, \text{uaugmented}]] \)

c. \( [[[\text{Class} –\text{singular}] [\text{Number} –\text{singular}, +\text{augmented}]] [D \text{usingular}, \text{uaugmented}]] \)

The content of D becomes (25) after valuation.

(25) a. \( [D \{ \}, \{–\text{augmented}\}] \)  <= (24a)

b. \( [D \{–\text{singular}\}, \{–\text{augmented}\}] \)  <= (24b)

c. \( [D \{–\text{singular}\}, \{+\text{augmented}\}] \)  <= (24c)

Only (25a) is realized as inverse, a correct result.

Other classes work in a similar way. Take the first person pronoun, which is inherently [+singular]. (26) gives rise to (27), which predicts correctly that this noun will exhibit the sii behavior.

(26) a. \( [[[\text{Class} +\text{singular}] [\text{Number} +\text{singular}, –\text{augmented}]] [D \text{usingular}, \text{uaugmented}]] \)

b. \( [[[\text{Class} +\text{singular}] [\text{Number} –\text{singular}, –\text{augmented}]] [D \text{usingular}, \text{uaugmented}]] \)

c. \( [[[\text{Class} +\text{singular}] [\text{Number} –\text{singular}, +\text{augmented}]] [D \text{usingular}, \text{uaugmented}]] \)

(27) a. \( [D \{+\text{singular}\}, \{–\text{augmented}\}] \)  <= (26a)

b. \( [D \{ \}, \{–\text{augmented}\}] \)  <= (26b)

c. \( [D \{ \}, \{+\text{augmented}\}] \)  <= (26c)

The function of [±group] is retained as before. The statements in (19) are brute-force instructions, overriding the general mechanism of valuation both in my account and in
Harbour’s. One might want to improve on (19), but that project would go far beyond the scope of this article.

(23) handles the rest of the classes in (11) correctly, as can be verified easily. I will leave the task to the reader.

Since the noun classes in Jemez are featurally identical to those in Kiowa minus the [+singular] Kiowa first person pronoun, the assumptions adopted so far predict that the pattern of inverse marking will be essentially the same, which is of course not the case. What is responsible for the difference between Jemez and Kiowa?

The clue is found when the lists in (11) and (12) are compared carefully. Recall Harbour’s (2011) observation about Jemez that when a class allows inverse marking, dual is always inverse-marked. The special status of dual is not limited to that point, however. Note that there is a systematic correspondence between Kiowa classes and their Jemez counterparts. When Kiowa has an $XY$ class, Jemez has an $XIY$ class, and vice versa. To see this correspondence, let us put noun classes in Kiowa and Jemez side by side.

(28) Noun classes in Kiowa and Jemez

<table>
<thead>
<tr>
<th>Kiowa class features</th>
<th>Jemez</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDP $\emptyset$</td>
<td>SIP</td>
</tr>
<tr>
<td>SII [+singular]</td>
<td>IIP</td>
</tr>
<tr>
<td>SII [–augmented]</td>
<td>IIP</td>
</tr>
<tr>
<td>SDI [–singular]</td>
<td>IIP</td>
</tr>
<tr>
<td>SDI [–singular, –augmented]</td>
<td>III</td>
</tr>
<tr>
<td>IDP [–singular]</td>
<td>IIIP</td>
</tr>
<tr>
<td>IDP [+augmented]</td>
<td>IIP</td>
</tr>
<tr>
<td>IDS [–singular, +group]</td>
<td>IIIS</td>
</tr>
<tr>
<td>SSS [–augmented, +group]</td>
<td>SSS</td>
</tr>
<tr>
<td>PPP [–augmented, –group]</td>
<td>PPP</td>
</tr>
</tbody>
</table>

The correspondence in question has no exception. Furthermore, it should also be pointed out that except for the SII class, which is absent in Jemez, Kiowa nouns never mark dual with inverse. The almost complete absence of inverse-marked dual in Kiowa is due to the fact that putting aside the SII class, constituted by the first person pronoun alone, and the [+augmented, ±group] classes, which are taken care of by the special provisions in (19), inherent class specification employs only negative values of [+singular] and ±augmented. Since dual is defined as [–singular, –augmented], Class and Number always share negative values in dual, producing no empty set in D after valuation.

So, if there is a way of converting the representation for dual into the one appropriate for inverse marking as defined in (23), we can account for the systematic correspondence between Kiowa and Jemez noun classes. And indeed, there is an independently motivated mechanism that can achieve our goal. Let us make some detour to review it.

Nevins (2011) discusses cases where the special dual form ceases to be used and is replaced by the plural form, as schematically shown in (29).

(29) Obliteration of the dual form in a certain environment

i. Environment X $\Rightarrow$ ii. Environment Y

| a. dual: A | a. dual: B |
| b. plural: B | b. plural: B |

He accounts for this extension of the plural form B to dual by means of impoverishment, which is posited in Distributed Morphology to delete features prior to Vocabulary Insertion, affecting the way Vocabulary items are inserted (Bonet 1991, Halle 1997, Harley 2008, and Noyer 1997, among others). He treats the impoverishment operation in question as
markedness-targeted, formulating it as follows:

(30) Impoverishment in environment Y
    Delete marked [–augmented] in the context of [–singular] in environment Y.

The crucial markedness statement is (31).

(31) Context-sensitive markedness statement
    In the context of [–singular], the marked value of [±augmented] is –.

The syncretism of dual and plural in environment Y in (29) is accounted for by the operation in (30), on the assumption that the relevant Vocabulary items are defined as in (32).

(32) a. A ⇔ [–singular, –augmented]
    b. B ⇔ [–singular]

After deletion of [–augmented] in environment Y, the Vocabulary item in (32a) becomes ineligible for insertion due to the Subset Principle in (33), the one in (32b) taking over instead.

(33) The Subset Principle (Halle 1997: 428)
    The phonological exponent of a Vocabulary item is inserted into a morpheme in the
    terminal string if the item matches all or a subset of the grammatical features specified
    in the terminal morpheme. Insertion does not take place if the Vocabulary item contains
    features not present in the morpheme. Where several Vocabulary items meet the
    conditions for insertion, the item matching the greatest number of features specified in
    the terminal morpheme must be chosen.

In environment X, where [–augmented] is not deleted, the item in (32a) has more matching features for dual than the one in (32b), and hence is chosen for insertion in dual. For concrete examples from Northern Sámi and Warlpiri manifesting the abstract pattern just illustrated, see Nevins’s article.

What is crucial for our purposes is the impoverishment operation in (30). Let us consider the representation of D for dual right after valuation, which is (34) for all the classes in Jemez and Kiowa except for the [+singular] SII pronoun in Kiowa and the two [+augmented, ±group] classes.

(34)   [D {–singular}, {–augmented}]

Suppose that impoverishment deletes αF, not the set whose member is αF. If (30) applies to D in Jemez, we get from (34):

(35)   [D {–singular}, { }] 

Since an empty set is created, inverse is expected, correctly. Thus, the difference between Jemez and Kiowa boils down to Jemez making use of the operation in (30).11

Significantly, the uniform characterization of inverse based on an additional operation of impoverishment in Jemez is made possible by the hypothesis that inverse is the exponent corresponding to the empty set as in (23). If [–F, +F] characterizes inverse as in Harbour’s

---

11 Demonstratives in Jemez pattern with the head noun in inverse marking, as can be seen from (13). This means that (30) applies to demonstratives as well.
original analysis, either [+singular] or [+augmented] must be added to (34) to turn the syntactically dual representation into the one appropriate for inverse. But there is no independently motivated mechanism that adds to (34) either [+singular] or [+augmented]. In this respect, the new theory of agreement plays a pivotal role in uncovering the fact that the nature of inverse is identical in Jemez and Kiowa. Note also that by adopting (23), the α-notation used in (15b) can be done away with. The uniform treatment of inverse in (23) is a definite improvement over Harbour’s analysis in this respect, too.

3.4. Verb Agreement

We have seen that a single uniform characterization of inverse suffices for Jemez and Kiowa, the difference in the number profile of noun classes with regard to dual being attributed to an operation of impoverishment deleting [–augmented] in Jemez. Since impoverishment applies after narrow syntax computation is over, the proposed analysis predicts that dual is not inverse-marked in narrow syntax in Jemez. This prediction is borne out by the behavior of verbal agreement. Let us consider relevant data.

Third person intransitive prefixes in Jemez are as follows:

(36) Third person intransitive prefixes in Jemez (based on Noyer 1997: 170–171)

<table>
<thead>
<tr>
<th>singular</th>
<th>dual</th>
<th>plural</th>
<th>inverse</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ø-</td>
<td>ɨl-</td>
<td>ɨl-</td>
<td>e-</td>
</tr>
</tbody>
</table>

Interestingly, apart from the SSS and PPP classes, the dual subject always triggers dual agreement, even though it is inverse-marked, as illustrated in (32) with an SII noun.

(37) a. Pá-nilší. (Yumitani 1998: 100)
    deer 3sg-fall.off.sg/du
    ‘A deer fell off.’

b. Pá-nilší.
    deer-inv 3du-fall.off.sg/du
    ‘Two deer fell off.’

c. Pá-nilší.
    deer-inv 3inv-fall.off.pl
    ‘Deer fell off.’

Harbour (2011) accommodates this fact by positing a special dual agreement marker corresponding to [–singular, –augmented]. Under his approach, the feature content of the clausal probe becomes identical to that of the D head after valuation. He claims that Pānini’s principle chooses the form for [–singular, –augmented] over the one for [αF, αG], which is less specific, attributing the use of the dual agreement marker for inverse-marked dual subjects in Jemez to the accidental existence of a special exponent for dual.12

12 If the Subset Principle in (33) is the technical rendition of Pānini’s principle, though, it is not obvious whether this result is guaranteed, since (33) only cares about the number of matching features. An additional clause needs to be added.

More seriously, it may not be correct to characterize both the nominal inverse marking and the verbal inverse agreement marker as [αF, αG] even under the set of assumptions adopted by Harbour. Inverse-marked nominals always contain [–singular, –augmented] at D under Harbour’s analysis, including duals, given that only the minus value of [±singular] or [±augmented] is used to define noun classes in Jemez, except for the [±augmented, ±group] classes. Then, the nominal inverse marking in Jemez should be characterized as [–singular, –augmented]. Furthermore, a more appropriate treatment of inverse verbal agreement under Harbour’s theory of agreement would be to posit the item
Under our new analysis, the persistent use of the dual agreement marker is simply due to the fact that the D head of dual subjects lacks the feature content corresponding to inverse marking in narrow syntax. The clausal probe is matched with the dual marking of D and is valued as such. Inverse marking of dual subjects is made possible by application to D (as well as to demonstratives) of the impoverishment operation that deletes \([-\text{augmented}]\) in the post-syntactic morphological computation, which does not have any impact on the working of the agreement operation in narrow syntax. Therefore, apart from this impoverishment operation, nothing special needs to be said about why inverse-marked dual subjects in Jemez do not trigger inverse agreement, unlike inverse-marked non-dual subjects.

But let us be more explicit about how valuation is executed at the clausal level. Take the sii noun in (37) again. Its inherent class specification is \([-\text{augmented}]\). In narrow syntax, only its plural form carries the feature content that induces inverse marking at D, which is valued as \([-\text{singular}, \{\}\] . The question is what feature specification the probe for clausal agreement ends up with when it agrees with \([-\text{singular}, \{\}\] . More generally, how does \{\} value the uninterpretable counterpart of the probe? I would like to suggest that a natural interpretation of the deletion algorithm in (21) dictates that no deletion be allowed. Given an empty set as the goal, it is impossible to define the feature whose value is opposite to that of the goal. This means that no number feature can be deleted from the probe. For (37), then, the probe for clausal agreement is valued as:

\[
\begin{align*}
(38)\ a. \ [+\text{singular}, \text{-augmented}] & = (37a) \\
\ b. \ [-\text{singular}, \text{-augmented}] & = (37b) \\
\ c. \ [-\text{singular}, +\text{augmented}, -\text{augmented}] & = (37c)
\end{align*}
\]

I posit the following agreement prefixes for third person intransitive subjects:

\[
(39)\ a. /\text{il}-/ \leftrightarrow [\text{-singular}, \text{-augmented}] \\
\ b. /\text{il}-/ \leftrightarrow [\text{-singular}, +\text{augmented}] \\
\ c. /\text{e}-/ \leftrightarrow [+\text{F}, -\text{F}]
\]

I attribute the zero marking for singular to the absence of the corresponding exponent. Note that the item in (39c) is correctly prevented from being inserted in the representation of dual in (38b). Only (39a) is eligible.

A different problem arises, however. Given the Vocabulary items in (39), it is not a trivial task to make sure that (39c) rather than (39a) or (39b) will be inserted into the probe valued as (38c). In terms of the matching features, (39a) and (39b) do not differ from (39c) in the case of (38c). What is going on?13

I would like to point out that the principle that ensures this result is cyclicity. Note that the correct result will be obtained if the most deeply embedded feature set, \{+\text{augmented}, -\text{augmented}\}, is picked out for Vocabulary Insertion in (37c) = (38c) . Embick (2010) claims that Vocabulary Insertion takes place cyclically. The analysis of Kiowa-Tanoan proposed in this article demonstrates that the domain of cyclic computation starts at the level of feature corresponding to \([-\text{singular}, \text{-augmented}, +\text{F}]\) rather than to \([\alpha\text{F}, \alpha\text{G}]\), since the non-dual inverse agreement marker always contains either \([\text{+singular}]\) or \([\text{+augmented}]\). The form for \([-\text{singular}, -\text{augmented}, +\text{F}]\) is simply ineligible for agreement with dual subjects. Conversely, the dual agreement marker is discarded for non-dual inverse by the straightforward application of (33).

13 Incidentally, the plural marker in Jemez cannot be assumed to be simply \[-\text{singular}\], unlike in the cases studied by Nevins (2011). Notice that \[-\text{singular}\] is not consumed by the inverse agreement marker in the case of (37c) = (38c). It is then unclear why the exponent for \[-\text{singular}\] is blocked from insertion in (37c) = (38c), if there is such an item available.
structure.

To see that if cyclicity is put aside, all the three items in (39) are eligible for (38c), consider a different agreement system in Jemez. Harbour (2011) notes syncretism of dual and inverse for what Yumitani (1998) calls intransitive benefactive prefixes, which are used in examples like those in (13), repeated below.

(13) a. Nı́́-t'aé  tê-êmê-Ø  nî́ i-l-k'á.
    (Yumitani 1998: 126)
    that-inv box-inv I 1sg-3inv-lie.sg/du
    ‘That box is mine.’
    ‘Those [two] boxes are mine.’
  b. Nî́ tê-êmê nî́ i-Ø-g'ó.
    that box I 1sg-Ø-lie.pl
    ‘Those [several] boxes are mine.’

The prefix consists of two pieces, as indicated in the gloss. The first part encodes the possessor, and the second the possessed. The alternation of the latter is characterized as /l/ for inverse versus zero for non-inverse by Yumitani (1998: 186), where inverse is understood to correspond to the actual marking on the noun, unlike in the account proposed in this article. Since the head noun in (13) belongs to the IIP class, which is inherently [–singular], the probe for the possessed is valued in narrow syntax as:

(40) a. [+singular, –singular], {–augmented} = (13a) with the singular reading
  b. {–singular}, {–augmented};] = (13a) with the dual reading
  c. {–singular}, [+augmented}] = (13b)

{+singular, –singular} in (40a) arises from agreement with { } on the goal. Recall that dual nominals behave as non-inverse in narrow syntax and are morphologically inverse-marked due to the operation of impoverishment.

I would like to suggest that /l/ is not the inverse agreement marker but the exponent for dual, as defined in (41).

(41)   /l/  ⇔ [–singular, –augmented]

The Subset Principle in (33) straightforwardly allows (41) to be inserted in the context of (40a) as well as (40b).

Essentially the same thing happens with SII nouns, as illustrated in (42).14

(42) a. Nɛ́-défi  nî́ i-Ø-k'á.
    (Yumitani 1998: 125-126)
    that chicken I 1sg-Ø-lie.sg/du
    ‘That chicken is mine.’
  b. Nî́-t'aé  défi-Ø  nî́ i-l-k'á.
    that-inv chicken-inv I 1sg-3inv-lie.sg/du
    ‘Those [two] chickens are mine.’
  c. Nî́-t'aé  défi-Ø  nî́ i-l-g'ó.
    that-inv chicken-inv I 1sg-3inv-lie.pl
    ‘Those [several] chickens are mine.’

The result of agreement at the clausal level is shown in (43), which is no different from (38).

14 Though the prefix in question is glossed as 3inv, it should be understood as in (41).
(43) a. [+singular], [+augmented] = (42a)
b. [–singular], [–augmented] = (42b)
c. [–singular], [+augmented, –augmented] = (42c)
(43c) as well as (43b) is qualified for being realized by (41). And this is exactly the source of the problem for intransitives discussed above, which calls for the cyclicity solution.

The difference between simple intransitive agreement in (37) and benefactive intransitive agreement in (42) in the behavior of the dual marker, then, comes from whether Vocabulary Insertion faces competition (37) or not (42). For benefactive intransitives, there is only one Vocabulary item encoding the number marking of the possessed noun. For simple intransitives, on the other hand, the item for dual must compete with another prefix (and loses in the case of non-dual inverse).

The verbal agreement in Kiowa requires no comments. The presence of { } on D values the probe as {+F, –F}, giving exactly the same results as Harbour’s account. In other words, our theory of agreement and Harbour’s do not differ empirically in the verbal domain, as far as the valued feature content of the clausal probe is concerned. The discussion in this section is therefore nothing more than a suggestion as to how Vocabulary items for verbal agreement should be handled in Jemez on the basis of the output of the narrow syntax computation, which does not differentiate between the two theories of agreement. This also means that the problem requiring the cyclicity solution, noted for Jemez above, arises for Kiowa as well, if Vocabulary items analogous to those in (39) are posited for Kiowa. My basic point in any case is that the relation between inverse marking on nominals and inverse agreement is made less transparent in Jemez by the operation of impoverishment in the nominal domain.

4. Conclusion

I have put forward a new theory of agreement that employs a simple algorithm of deletion for the purposes of valuation, building on Harbour’s (2007, 2011) recent proposal concerning uninterpretable features, which is promising as part of the principled account of the evolutionary foundation of the Agree operation. The new theory makes it possible to provide a novel, uniform characterization of inverse in Jemez and Kiowa. It also sheds new light on how clausal agreement works in Jemez. Thus, it is empirically attractive, too.

The discussion of verbal agreement leads to the discovery that the cyclic nature of linguistic computation can be observed even at the level of feature structure. Though not noted by Harbour, his conception of uninterpretable features suggests some degree of hierarchical organization in the feature matrix. The cyclic Vocabulary Insertion made use of in this article rests on that structural property. It is an interesting topic for future research to explore whether further structure can be found in the organization of φ-features.

References


These conclusions entail that a more appropriate alternative for the Jemez verbal agreement marker under Harbour’s account pointed out in note 12 can in principle be adopted for our account of Jemez. Then, one needs to ask whether there are any reasons for preferring (39c) to the possibility raised in note 12. One major consideration is that (39c) uses fewer features for the characterization of the inverse agreement marker than [–singular, –augmented, +F]. Another is that (39c) transparently reflects the result of agreement with an empty set.


SIMPLEST MERGE GENERATES SET INTERSECTION:
IMPLICATIONS FOR COMPLEMENTIZER ‘TRACE’ EXPLANATION*

Samuel D. Epstein¹, Hisatsugu Kitahara², and T. Daniel Seely³
University of Michigan¹, Keio University², and Eastern Michigan University³

1. Introduction

A unique aspect of the analysis presented in this paper is that even though we propose Merge in its simplest, strictly binary form, it is nonetheless the case that a category can have more than one sister (i.e. more than one element that it was merged with). The multiple sisters of a given category arise at different derivational points, generating a “two-peaked” structure, or more formally, intersecting sets, neither of which is a term of the other, in the set-theoretic notation of bare phrase structure (Chomsky 1995a). We assume an analysis with a very general, independently motivated deletion algorithm, which retains only the highest copy of multiple copies of an internally merged element, for sensorimotor SM systems. This is presumably an application of the independently necessary overarching principle of Minimal Computation (categories or features must be retained in an optimal fashion) entailing that copies cannot have additional phonological features above and beyond the single set of lexical features of the mover. Conversely, we also assume the principle of recoverability of deletion (categories or features must not be deleted in a random fashion). We then argue that our proposed simplest formulation of Merge (generating set intersection when applied “countercyclically”) coupled with the very general laws of Minimal Computation (informally forcing features to be interpreted at most once) and Recoverability (informally forcing features to be interpreted at least once) allows us to deduce the core complementizer-trace phenomena with no further mechanisms required. Complementizer-trace effects turn out to be an immediate consequence of these arguably quite natural and general principles, having no ad hoc language-specific, construction-specific or operation-specific motivation regarding complementizer-trace phenomena. We then address the cross-linguistic variation of complementizer-trace effects and the apparent problem of undergeneration that our analysis faces given that it deduces complementizer-trace effects from deep, i.e. unparameterizable principles.

2. Simplest Merge and its Inevitable Consequence for the Derivation of “Countercyclic Movement”

In this subsection, we provide a brief overview of the development of the simplest conception of Merge from X-bar theory through bare phrase structure in the sense of Chomsky (1995a). The goal is to factor out of the Merge operation any property that can itself be deduced from deeper principles, leaving the optimal structure building operation, what we refer to below as simplest Merge.

* We would like to thank especially Noam Chomsky for his interest in our work, and also, Miki Obata and Mamoru Saito. All errors are, of course, our own. Part of the material in this paper was presented at Glow in Asia IX, hosted by Mie University (September 5, 2012). We are grateful to the participants for clarifying remarks.
2.1. Brief Overview of X-bar Theory in Early Minimalism

X-bar theory sought to eliminate phrase structure rules, leaving only the general X-bar-format as part of UG. Determining that format, and avoiding construction-specific and language-specific phrase structure rules, was the central research goal of subsequent work. In early minimalism, X-bar theory was still given, with specific stipulated properties. The outputs of the applications of structure-building operations – binary and singulary Generalized Transformation GT – were then assumed to be constrained by X-bar theory, by definition (see Chomsky 1993):

(1) a. binary GT
   (i) takes K and K’, where K’ is not a phrase marker within K
   (ii) adds Δ external to K
   (iii) substitutes K’ for Δ, forming K*, which must satisfy X-bar theory

b. singulary GT
   (i) takes K and K’, where K’ is a phrase marker within K
   (ii) adds Δ external to K
   (iii) substitutes K’ for Δ, forming K*, which must satisfy X-bar theory

Basically, binary GT takes two separate syntactic objects and combines them into a single object, which as we’ll see in a moment is the ‘ancestor’ of External Merge. Singulary GT is the precursor of the more recent Internal Merge, where one of the objects being joined together is contained within the other. In effect, X-bar theory, together with its stipulated (or axiomatic) properties (endocentricity, head-to-complement, and spec-head relation) was taken to be a UG filter on transformational output representations.

2.2. Subjecting X-bar Theory to a Minimalist Critique

Under the strong minimalist thesis, however, X-bar theory was not exempt from explanatory scrutiny; it was asked why X-bar theory seems to hold, as opposed to an infinite number of formally definable alternative phrase structure systems. By adherence to Minimalist method, this question prompts the following question: How “should” phrase structures be generated under minimalist (ideal, simplest) assumptions? Chomsky’s (1995a:396) answer was:

Given the numeration N, C_{HL} may select an item from N (reducing its index) or perform some permitted operation on the structure it has already formed. One such operation is necessary on conceptual grounds alone: an operation that forms larger units out of those already constructed, call it Merge. Applied to two objects α and β, Merge forms the new object γ. What is γ? γ must be constituted somehow from the two items α and β; ... The simplest object constructed from α and β is the set {α, β}, so we take γ to be at least this set, where α and β are constituents of γ. Does that suffice? Output conditions dictate otherwise; thus verbal and nominal elements are interpreted differently at LF and behave differently in the phonological component ... γ must therefore at least (and we assume at most) be of the form {δ, {α, β}}, where δ identifies the relevant properties of γ, call δ the label of γ.

Merge was introduced as an operation (the central structure building operation of the narrow syntax NS), necessary on conceptual grounds alone, and the simplest object γ constructed from α and β by Merge was taken to be the set {α, β}. Chomsky (1995a) assumed the set {α,
β} was too simple; it was assumed that empirical adequacy demanded some departure from the simplest assumption (the standard scientific tension between explanation and “empirical coverage”); that is, the set must be labeled as in e.g. \{δ, \{α, β\}\}, where δ identifies the relevant properties of γ.

Given that an output of Merge is a labeled set γ=\{δ, \{α, β\}\}, Chomsky (1995a:397-398) asked what exactly the label of γ is:

If constituents α, β of γ have been formed in the course of computation, one of the two must project, say α. At the LF interface, γ (if maximal) is interpreted as a phrase of the type α (e.g. a nominal phrase if its head κ is nominal), and it behaves in the same manner in the course of computation. It is natural, then, to take the label of γ to be not α itself but rather κ, the head of the constituent that projects, a decision that also leads to technical simplification. Assuming so, we take γ = \{κ, \{α, β\}\}, where κ is the head of α and its label as well.

Under this definition, the label of γ is the head of one of its constituents. If α projects, then the object γ constructed from α and β by Merge is \{H(α), \{α, β\}\}, where H(α) is the head of α (see also Chomsky 1995b). Additionally, the notion “term” is defined as follows: (i) K is a term of K; and (ii) if L is a term of K, then the members of the members of L are terms of K (Chomsky 1995a:399).

Chomsky (1995a,b) did not discuss exactly how Merge operates to form such labeled sets, but one way is to formulate Merge as an operation consisting of the following two steps:

\begin{enumerate}
\item Applied to α and β, where neither α nor β is a term of the other, Merge
  \begin{enumerate}
  \item takes α and β, forming \{α, β\}, and
  \item takes H(α) and \{α, β\}, forming \{H(α), \{α, β\}\}.
  \end{enumerate}
\item Applied to α and β, where α is a term of β, Merge
  \begin{enumerate}
  \item takes α and β, forming \{α, β\}, and
  \item takes H(β) and \{α, β\}, forming \{H(β), \{α, β\}\}.
  \end{enumerate}
\end{enumerate}

In (2a,b), the second step in effect labels the simplest object \{α, β\} constructed by the first step. A question is whether we can eliminate or derive the empirically desirable aspects of this second step – i.e. whether we can predict or explain by general principles what the label of any α, β pair will be (see Chomsky 2000:133). If the answer is positive, Merge can be formulated in the simplest form: Merge(α, β)\Rightarrow\{α, β\}, with the label eliminated from the representational notation and instead simply identified as H(α) or H(β), as in Chomsky (2012).

2.3. On the Complexity of “Countercyclic” Covert Movement

In addition to the labeling algorithm, Chomsky (1995b:254) noted an additional complexity, one concerning covert movement:

The computational system C_{HL} is based on two operations, Merge and Move. We have

\[1\] See Collins (2002) and Seely (2006) for discussion of the idea that such labels (and label projection) can be eliminated entirely from the grammar.

\[2\] Chomsky (1995a,b) assumes that either α or β may project (in principle), but if the wrong choice is made, deviance would result.
assumed further that Merge always applies in the simplest possible form: at the root. What about Move? The simplest case again is application at the root: if the derivation has reached the stage $\Sigma$, then Move selects $\alpha$ and target $\Sigma$, forming $\{\gamma, \{\alpha, \Sigma\}\}$. But covert movement typically embeds $\alpha$ and therefore takes a more complex form: given $\Sigma$, select $K$ within $\Sigma$ and raise $\alpha$ to target $K$, forming $\{\gamma, \{\alpha, K\}\}$, which substitutes for $K$ in $\Sigma$.

We can formally represent this additional complexity by adding a third step to the formulation of Merge:  

(3) Applied to $\alpha$ and $\beta$ within $\Sigma$, where $\alpha$ is a term of $\beta$, Merge
   (i) takes $\alpha$ and $\beta$, forming $\{\alpha, \beta\}$, and
   (ii) takes $H(\beta)$ and $\{\alpha, \beta\}$, forming $\{H(\beta), \{\alpha, \beta\}\}$, and
   (iii) replaces $\beta$ in $\Sigma$ by $\{H(\beta), \{\alpha, \beta\}\}$

Again, a question is whether we can eliminate this third step (countercyclic IM replacement, or ‘substitution’ to use Chomsky’s (1995b:254) terminology), along with the second step (labeling), and thereby derive an empirically adequate, simplest formulation of Merge. If the answer is positive, Merge can be formulated in the simplest form: Merge $(\alpha, \beta) \Rightarrow \{\alpha, \beta\}$.

2.4. An Inevitable Consequence of Simplest Merge

In a series of recent papers, Epstein, Kitahara, and Seely (hereon EKS) have proposed precisely this form of simplest Merge. Unlike Chomsky’s Merge, which allows “countercyclic” replacement as in (3iii), simplest Merge cannot perform replacement. EKS have explored its effects (see EKS 2012, Kitahara 2011) including the inevitable consequences of simplest Merge. Let’s review. Consider the following sentence:

(4) Bill ate rice.

At some point in the derivation of (4), Merge takes the phase-head C and merges it with TP, yielding (5):

(5) $\{C, \{T, \{Bill, \{v, \{ate, rice\}\}\}\}\}\}

At this point, C transmits unvalued phi to T. T, then, functioning as a phi-probe, locates the goal Bill (bearing lexically valued phi and unvalued Case), and Agree between T and Bill applies, valuing phi on T and Case on Bill. The standard assumption (see Chomsky 2007, 2008) is that as these features get valued, Bill is raised to the so-called Spec(ifier) of T (such raising required by some residue of EPP), yielding (6) (where we use the notion “Spec(ifier)” only for expository purposes):

(6) $\{C, \{Bill, \{T, \{Bill, \{v, \{ate, rice\}\}\}\}\}\}\}$

3 See Groat and O’Neil (1996) for an alternative approach to covert movement whereby covert movement applies cyclically in the syntax, but the chain-tail is pronounced, as if, in the eyes of PF, the movement has not applied (= the definition of “covert” movement).

4 In section 4, we suggest that simplest Merge is not only conceptually desirable, but empirically adequate, and that at least with respect to that-trace phenomena, notoriously resistant to explanation, the empirical coverage/explanatory power of the theory is in fact concomitantly increased.
This “countercyclic” application of Merge (which, recall is not simplest Merge) mapping (5) to (6) executes a form of replacement (= (3iii)) since Merge, not applying at the root, “infixes” *Bill* into SpecTP. In set-theoretic terms, “countercyclic” Merge of *Bill* to the then specless $T_1 = \{T, \{\text{Bill}, \{v, \{\text{ate, rice}\}\}\}\}$ yields a new specful syntactic object, namely $T_2 = \{\text{Bill}, \{T, \{\text{Bill}, \{v, \{\text{ate, rice}\}\}\}\}\}$. But recall, this creation of SpecTP is not at the root, since $C$ was necessarily already externally merged with the specless $TP = T_1$ (as in (5)). Therefore, $T_2$ (with *Bill* as ‘newly appointed’ SpecTP) replaces specless $T_1$ in $\{C, T_1\}$, removing $T_1$ from $\{C, T_1\}$ (destroying a relation/set) and merging $T_2$ as the newly appointed “sister” to $C$. Such replacement – $T_2$ replaces $T_1$, and $T_1$ “vanishes” – yields a new CP (= $\{C, T_2\}$). It is important to note that it is this “new” CP (= $\{C, T_2\}$), and not the “old” CP (= $\{C, T_1\}$), that enters into further derivational processes, i.e. the specless $TP = T_1$ has disappeared from the continuing derivation.

EKS (2012) argue that, given the strong minimalist thesis, NS should contain only the simplest structure-building operation, namely $\text{Merge}(X, Y) \Rightarrow \{X, Y\}$. EKS then point out that Merge, defined in this simplest form (= (3i) only) cannot replace existing categories (see Freidin 1999 for important earlier discussion of replacement as a non-primitive operation). Thus, it must be the case that NS equipped only with simplest Merge cannot map (5) (= $\{C, T_1\}$) to (6) (= $\{C, T_2\}$) in the manner just discussed above. But then, what kind of object does “counter-cyclic” application of Merge (e.g. raising to specTP) create?

It is important to note that EKS adopts the idea of Chomsky (2007, 2008) that given a lexical array (which itself must contain a phase head $C$ or $v$), all instances of EM must take place before any instance of IM. This is motivated on grounds of efficiency, specifically: EM involves no search (there is immediate access to lexical items within an array), while IM (and other operations of the NS such as Agree) involve search (into an already-existing syntactic object). Thus, it follows that EM must be exhausted before any other operation can apply. In (5), then, the NS builds up to CP via EM, and only then can IM take place, creating (6)—if replacement were allowed. Thus, the nature of “countercyclic” IM is crucial in that it is in effect “forced” under efficiency. Since, by hypothesis, $C$ is merged to TP before raising of the External Argument to SpecTP, it follows that when such raising does take place, it is “countercyclic” in the sense specified above.

EKS suggest that the “countercyclic” application of Merge cannot execute the complex operation “replacement” but rather necessarily results – speaking informally now – in a “two-peaked” or “doubly-rooted” representation, informally represented in (7) (where indices appear only for expository purposes, and linear order is irrelevant):

---

5 Chomsky (2007, 2008) and EKS (2012) argue, and in fact attempt to deduce, that all instances of EM must take place before any instance of IM, and hence $C$ must be merged to TP before any instance of IM (involving TP) can occur.
In set-theoretic terms, first we created the set $CP = \{C, T_1\}$ since $C$ was merged with specless $T_1$. Then, the "countercyclic" application of simplest Merge (incapable of replacement, by our assumptions) merges $Bill$ to $T_1$ forming another set, namely $T_2 = \{Bill, T_1\}$. This results in two distinct but intersecting set-theoretic syntactic objects SOs, which happen to have $T_1 (= \{T, vP\})$ as their "shared" element. That is, as a result of this "countercyclic" Merge (of $Bill$ to $T_1$), the single "workspace" of NS contains two intersecting set-theoretic SOs, $CP = \{C, T_1\}$ and $T_2 = \{Bill, T_1\}$, as shown in (8a,b), where neither is a term of the other.$^6$

(8) a. $\{C, \{T, \{Bill, \{v, \{ate, rice\}\}\}\}\}\}$
   b. $\{Bill, \{T, \{Bill, \{v, \{ate, rice\}\}\}\}\}\}$

Interestingly, notice that although Merge is the simplest possible operation, crucially binary and incapable of replacement thus creating only two-membered sets, it is nonetheless the case that $T_1 (= \{T, \{Bill, \{v, \{ate, rice\}\}\}\})$ has two different sisters. That is, simplest Merge generates intersecting sets: both $C$ and ("later") $Bill$ were merged with $T_1 (= \{T, \{Bill, \{v, \{ate, rice\}\}\}\})$ at different derivational points; hence, $C$ and $Bill$ are sisters of $T_1$.

Given this result, EKS (2012) further argue that, under the law of semantic composition, these two intersecting set-theoretic SOs, each functioning as a root, would not yield a single semantic value if they were sent to the semantic component together, "as one". Thus, there must be some way to decompose this "two-peaked" or "doubly-rooted" representation in the "workspace" of NS, prior to the semantic component. One possible way out of this situation, outlined by EKS (2012), is to send the two intersecting SOs to the semantic component separately. Specifically, EKS propose that Transfer dissolves this intersecting situation by removing one intersecting set from the "workspace" of NS. In effect, one "peak" (i.e. one of the two intersecting sets) must be sent to the semantic component by Transfer at this point of the derivation. Under this proposal, cyclic Transfer of TP once CP is built is deduced (given that the phase-edge must be left for the derivation to continue). Suppose that (i) the "workspace" contains (8a) (= \{C, T_1\}) and (8b) (= \{Bill, T_1\}), and (ii) cyclic Transfer removes (8b) (= \{Bill, T_1\}). Then, the phase-edge, namely \{C, ---\}, where "---" is transferred material, in effect is left for subsequent operations. Note, if $vP$, in addition to CP, is a phase, then cyclic Transfer of VP at $vP$ also follows.

Summarizing, the "countercyclic" application of Merge (raising Subject to SpecTP after the merger of C and specless TP) cannot map \{C, \{T, vP\}\} to \{C, \{Subject, \{T, vP\}\}\}. Instead, it necessarily forms \{Subject, \{T, vP\}\}, which intersects with \{C, \{T, vP\}\}, and neither \{Subject, \{T, vP\}\} nor \{C, \{T, vP\}\} is a term of the other. This "two-peaked tree" (or these intersecting sets) thus arises as an inevitable consequence of simplest Merge when

---

$^6$ (8a,b) is the single representation of the output of "countercyclic" subject raising. That is, there aren’t two "separate workspaces" denoted by (8a,b), but rather a single representation of intersecting sets, analogous to the single phrase structure representation in (7), which is correspondingly not interpreted as two "separate" trees.
applied “counter-cyclically.” In effect, Merge can create (but not destroy) syntactic relations. It naturally guarantees that all established syntactic relations (among existing categories) remain unaltered in NS.

3. **Simplification by Deletion**

In NS, all syntactic relations, each established by Merge, remain unaltered, but when transferred to SM, a set-theoretic SO will be simplified by deletion. Although an SO that enters the phonological component will contain all relevant lexical features and all relevant syntactic structures (created in the course of its derivation), by SM what will remain of this SO is just a “pure” phonological representation; all but the phonological features will be deleted from it. What is particularly relevant for present concerns is the case of movement, i.e. cases where there are multiple copies of a mover. Here too, there will be simplification by deletion, specifically, only one copy of a moved element is present at SM, all other copies are deleted. In this section, we review the relevant assumptions concerning deletion in the phonological component, setting the stage for our analysis of that-trace effects.

One of the underlying and enduring assumptions in generative grammar is the principle of recoverability (Recoverability). This principle states that each SM or conceptual-intentional CI interpretable feature borne by a lexical item must be present at the relevant interface. Thus, no CI or SM interpretable feature of a lexical item can simply vanish in the course of a derivation. Thus, the lexical item *cot* with its phonological features (associated with /kat/) cannot be realized in SM as just /ka/ with no sign of /t/. Likewise if the overt complementizer *C* 

\[ \text{that} \]

enters into a derivation, then its phonological features, represented (informally) as /that/, must be present in the phonological representation; these phonological features can’t be lost or vanish in the course of the derivation. So, Recoverability guarantees the presence (at an interface) of relevant CI and SM interpretable features. But, while by hypothesis empirically necessary (given Recoverability), mere presence at the interface is not empirically sufficient.

Another basic assumption, assigned more prominence in recent minimalist literature, is the principle of minimal computation (Minimal Computation). This principle states that an interface-interpretable lexical feature, which must be present (at least once) at the interface given Recoverability, cannot be present more than once. Chomsky (2012) notes that “universally in language, only the structurally prominent copy is pronounced,” which, he suggests “follows from another application of the third factor principle of Minimal Computation: Pronounce as little as possible.” So, for example, if a category bearing phonological features, e.g. *Bill*, undergoes movement, then there will be two copies of *Bill* in the output, but only the structurally prominent (universally the highest) copy of *Bill* will be phonologically realized as /Bill/. Consider (9):

\[
(9) \quad \{\text{Bill, } \{T, \{\text{Bill, } \{v, \text{left}\}\}\}\}\}
\]

In (9), assuming that *Bill* moves just from SpecP to SpecTP, there would be two copies of *Bill*, only the highest of which is “pronounced,” hence at SM it’s /Bill left/ and not /Bill Bill left/. The relative height of such copies is calculated by the positions of their occurrences, where, following Chomsky (1995b), an occurrence of X is defined as a category to which X is merged, i.e. its derivational sister. Thus, in (9) there is one *Bill* (i.e. just one category, *Bill*, entered the derivation via a lexical array) with two occurrences determined by *Bill*’s sisters, namely, the two occurrences \{v, left\} and \{T, \{Bill, \{v, left\}\}\}. The latter occurrence is higher than the former occurrence since the latter occurrence contains the former occurrence as its term, hence *Bill* in SpecTP is the (only) copy that is phonologically realized. Under this occurrence-based calculation, the position of each copy of X is uniquely determinable by
reference to its merged sister, namely its occurrence. As the derivation proceeds, the merged sisters of the lower occurrences of X are all deleted, and only the highest copy of X will remain for SM.

To summarize, Recoverability requires that each CI and SM interpretable feature of a lexical item is present (hence present at least once) in the relevant interface representation. Minimal Computation requires that interpretable features are present at most once. Finally, the very general and completely independently necessary Full Interpretation guarantees that interface interpretation actually takes place. That is, the features that are (guaranteed to be) present in an interface representation can’t simply be ignored, but must be implemented by the interface. Putting this all together, independently of anything having to do with that-trace, the system determines that although the representation underlying, say, *Bill was arrested* involves [Bill was arrested Bill], the object *Bill* is phonologically realized once, but only once.

4. Complementizer-Trace Effects: A New Analysis

Adopting simplest Merge, along with its inevitable consequence of set-intersected representation (“doubly-rooted” tree), we argue that complementizer-trace effects, or at least the core cases discussed in the context of that-trace (see Abe 2011, Browning 1996, Chomsky 1981, 1986, Chomsky and Lasnik 1977, Ishii 2004,Lasnik and Saito 1984, 1992, Perlmutter 1971, Rizzi and Shlonsky 2007, Sobin 1987, 2002, among many others) are deducibly excluded, based on simplest Merge (as detailed above) and very general, deep principles, specifically: Recoverability and Minimal Computation. It is important to stress again that we are not adding any new principles or “technical” mechanisms of any sort, rather we are appealing to operations (simplest Merge) and deep principles (Recoverability and Minimal Computation) that are independently necessary and well-supported empirically. Our goal is to reveal that core complementizer-trace effects fall out as a natural consequence of components of the system that are “virtually conceptually necessary” or (seemingly) simplest – e.g. Merge as defined in (3i) is by hypothesis simpler than Merge defined as (3i, and ii, and iii).

4.1. Core Cases

Consider the following well-known contrast between (10) and (11):

(10) Who do you think left?

(11) * Who do you think that left?

Let’s consider the derivation of (10), key parts of the derivation of which are represented informally as in (12), with $C_2$ the embedded CP:
who do you think:

This tree representation informally depicts the “embedded” clause, which of course, at the “time” of its generation, has not yet become embedded. Recall that there is exactly one lexical item who associated with this derivation (who is selected exactly once from the lexicon) and this who has (at least) three copies in the embedded CP, the copies represented as who3, who2, who1 (where the indices are used just for expository convenience). The lexical item who originated in SpecvP, and it was “cyclically” merged with C1 yielding C2 and then it was “countercyclically” merged with T1 yielding T2. The creation of T2 is precisely the “two-peaked” situation we detailed above. In the course of this derivation, T1 = \{T, vP\} ends up having two different sisters, namely, Cnull, which T1 was cyclically merged with, and who2, which T1 was “countercyclically” merged with. Now, we know that the string of copies of who are not all pronounced (the final SM implementation of the derivation is certainly not /who do you think who who left/). Thus, the merged sisters of the lower occurrences of who are all deleted in accord with Minimal Computation.

To illustrate in more technical detail, the (relevant part of the) derivation “starts” with the lower CP phase:

(13) \{C, \{T, \{who1, \{v, {\text{left}}\}\}\}\}\}

Since the option of applying EM is now exhausted, phi feature (C to T) transmission and (then) Valuation (of phi of T and Case of who) are carried out. And, crucially for present concerns, there is IM of who to TP (to create SpecTP), and there is also IM of who to CP (to create SpecCP). These simultaneous applications of IM map (13) to both (14) and (15), (where (14) and (15) are informally represented in (12)):

(14) \{who2, \{T, \{who1, \{v, {\text{left}}\}\}\}\}\}

(15) \{who3, \{C, \{T, \{who1, \{v, {\text{left}}\}\}\}\}\}\}

At this point, there are three occurrences (i.e. derivational sisters) of who:
(16) a. \{v, \{left\}\}  
who was merged to \{v, \{left\}\},  
yielding \{who_1, \{v, \{left\}\}\}

b. \{T, \{who_1, \{v, \{left\}\}\}\}  
who was merged to \{T, \{who_1, \{v, \{left\}\}\}\},  
yielding \{who_2 \{T, \{who_1, \{v, \{left\}\}\}\}\}

c. \{C, \{T, \{who_1, \{v, \{left\}\}\}\}\}  
who was merged to \{C, \{T, \{who_1, \{v, \{left\}\}\}\}\},  
yielding \{who_3 \{C \{T, \{who_1, \{v, \{left\}\}\}\}\\}

Notice that at each derivational step in (16), in fact throughout any derivation, only simple, binary Merge is employed, and sisterhood is characterized derivationally in the simplest way, as “merged with” i.e., X and Y are sisters if and only if Merge(X, Y) applied (see Epstein, Groat, Kawashima and Kitahara 1998). But, interestingly, the object \{T, \{who, \{v, \{left\}\}\}\}, an occurrence of who, has two (derivational) sisters: namely, the category C (see (15)) and (what is now a copy of) who (see (14)). Given Minimal Computation, the single lexical item who, bearing one and only one set of phonological features that entered the derivation, must be pronounced once and only once, at the highest copy position. But at this point we must consider with some care just how Minimal Computation is to be stated.

The intuitive idea, as we’ve seen, is that only one (the highest) copy of a moved element is realized in the phonological component (only the highest is phonologically implemented). And what this entails is that all other copies determined by the lower occurrences of the moved element must be deleted, and such deletion takes place phase-by-phase. Now we know that in the mapping from NS to SM, there is massive deletion of all but the purely phonological (i.e. SM-interpretable) features; given the strong minimalist thesis, all and only SM interpretable features are present in the representation that exits the phonological component. Thus, all syntactic structures, all semantic features, and all but the highest copy of a moved element must be deleted (i.e. removed) from the SO that undergoes this mapping.

With this in mind, consider again (14) and (15). Assuming that every copy of who represented in (14) and (15) is determined to be the sister of a lower occurrence of who, it must be deleted. Recall, occurrences of X are the (derivational) sisters of X. So, the general deletion algorithm looks at \{T, \{who_1, \{v, \{left\}\}\}\} and \{v, \{left\}\} (since they are lower occurrences of who) and says: “Delete all their sisters, i.e. delete the who copies which are the sisters of \{T, \{who_1, \{v, \{left\}\}\}\} and of \{v, \{left\}\}. Thus, the operation deletes all the lower copies of who from (14) and (15). But that’s not the end of the story.

As we pointed out above, for the EKS analysis, even though Merge is itself simple(st) and binary, with “countercyclical” IM creating SpecTP, the object \{T, \{who_1, \{v, \{left\}\}\}\} can have two derivational sisters. In short, sister is not uniquely defined in that sisters of {T, {who, {v, {left}}}} and {v, {left}} (since they are lower occurrences of who) and says: “Delete all their sisters, i.e. delete the who copies which are the sisters of {T, {who, {v, {left}}}} and of {v, {left}}. Thus, the operation deletes all the lower copies of who from (14) and (15). But that’s not the end of the story.

As we pointed out above, for the EKS analysis, even though Merge is itself simple(st) and binary, with “countercyclical” IM creating SpecTP, the object \{T, \{who_1, \{v, \{left\}\}\}\} can have two derivational sisters. In short, sister is not uniquely defined in that sisters of {T, {who, {v, {left}}}} and {v, {left}} (since they are lower occurrences of who) and says: “Delete all their sisters, i.e. delete the who copies which are the sisters of {T, {who, {v, {left}}}} and of {v, {left}}. Thus, the operation deletes all the lower copies of who from (14) and (15). But that’s not the end of the story.
(a subject “trace”), it necessarily also deletes the local (and only the local) C – since the two are in the exact same position i.e. sister to \{T, vP\}. Overall, that-trace effects with subject (not object or adjunct) extraction follow, i.e. we can’t “say” who do you think that left for the same reason that we can’t say Bill was arrested Bill, in both cases we have “pronounced” too much in violation of Minimal Computation.

Returning to the that-trace contrast, exhibited between (10) and (11), it now follows immediately. Recall that, in the “doubly-rooted” tree or the corresponding intersecting set representation (12), C\_null counts as a merged sister of \{T, vP\}. Now, if the two lower copies of who get deleted, the phonological features of who are in principle recoverable, because the highest copy of who (in the edge of the lower CP phase at this point of the derivation) still remains. And since C\_null bears no phonological features, deleting the phonological features of C\_null is vacuous, with no problem resulting from this. Thus, the deletion of the two copies of who and C\_null does not violate Recoverability in the case of (10). Similarly, in the derivation of (11), the merged sisters of the lower occurrences of who are all deleted in accord with Minimal Computation. However, consider the corresponding “doubly-rooted” tree of (11), in which the complementizer bears phonological features, e.g. /that/, as given in (17):

(17) * who\_4 do you think:

There are (at least) two lower occurrences of who in the embedded CP: \{T, vP\} and \{v, VP\}. The merged sisters of these occurrences include two copies of who and C\_that. Now, if the two copies of who get deleted, the phonological features of who are in principle recoverable, because the highest copy of who (in the edge of the lower CP phase at this point of the derivation) still remains. So, the deletion of the two copies of who does not violate Recoverability. But what about the deletion of C\_that? Recall again, under the “two-peaked” or intersecting situation, represented in (17), C\_that counts as a merged sister of \{T, vP\}, and if the deletion rule must (indiscriminately) apply to all the merged sisters of the lower occurrences of who (understood to be the simplest application of this deletion algorithm), then C\_that must undergo deletion, together with who\_2 (= SpecTP) and who\_1 (= SpecvP). In other words, the deletion of C\_that is an inevitable consequence of the demand, imposed by Recoverability (“not too little”) and Minimal Computation (“not too much”). But C\_that does bear phonological features (and it is the sole copy), so the deletion of C\_that violates Recoverability. That is, that with a single set of phonological features never underwent SM interpretation of those phonological features. Thus, the only derivation that survives when performing subject extraction, is the one in which C\_null is (randomly) chosen (and this is without look-ahead, of course).

Notice, if we analyze if as C\_if, then the contrast between (18) and (19) follows in the same way:
(18) ? how many cars did he wonder [\(t_{\text{wh}} C_{\text{if}} [\text{the mechanics fixed } t_{\text{wh}}]\)]

(19) * how many mechanics did he wonder [\(t_{\text{wh}} C_{\text{if}} [t_{\text{wh}} \text{ fixed the cars}]\)]

Observing this contrast, Chomsky (2012) notes that what (19) expresses is “a fine thought, but it has to be expressed by some circumlocution.” He speculates that there is something about language design, which poses a barrier to communication. We suggest that Recoverability and Minimal Computation constitute just such barriers.

4.2. Related Cases

As we have seen in the derivation of (11), if \(C_{\text{that}}\) is chosen, the derivation is in trouble because the deletion rule (applying to SpecTP) will “unrecoverably” delete \(\text{that}\) occupying the same position as SpecTP which itself undergoes deletion. So, if the deletion rule does not apply to SpecTP, e.g., if SpecTP is phonologically realized, then \(C_{\text{that}}\) will not be in trouble, and it will be phonologically realized as /that/. This prediction is confirmed. Consider (20), in which SpecTP is phonologically realized as /John/, and (21), in which SpecTP is phonologically realized as /there/ (Safir 1985, Chomsky 1991, 1995b:158):

(20) Who do you think \([CP \ t_{\text{wh}} C_{\text{that}} [TP \text{ John likes } t_{\text{wh}}]]\]

(21) How many men did John say \([CP \ t_{\text{how}} C_{\text{that}} [TP \text{ there were } t_{\text{how}} \text{ in the room}]]\]

In each derivation, a \(\text{wh}\) -element moves directly to (embedded) SpecCP (i.e. there is no movement of \(\text{wh}\) -element to SpecTP), so the deletion rule does not apply to SpecTP; therefore, as shown above, it does not apply to the local \(C\). Instead, SpecTP is phonologically realized, and \(C_{\text{that}}\) is phonologically realized as /that/. There is no obligatory deletion of \(C\) in such cases, and Recoverability is thereby satisfied. Similarly, this analysis captures the non-deviant status of (22):

(22) Why do you think \([I_{\text{wh}} C_{\text{that}} [\text{John likes Mary } t_{\text{wh}}]]\]

Again, in constructing the embedded CP, the \(\text{wh}\) -element, \(\text{why}\) in this case, is never compelled to move to SpecTP, instead moving directly to SpecCP, and hence is never a derivational sister to \(\{T, vP\}\) and hence \(\text{that}\) escapes (obligatory but Recoverability-violating) deletion. In fact, our analysis captures the notorious subject/adjunct asymmetry in \(\gamma\) -marking (see Huang 1982, Lasnik and Saito 1984, 1992) but without invoking any SS versus LF asymmetry in the targets (arguments versus nonarguments) of the \(\gamma\) -marking algorithm.\(^7\) All such descriptive technicalia are prohibited under minimalist analysis rightly raising the bar on what can count as an explanation versus an unilluminating formal/technical re-description often more “complex” than the data described (as discussed by Chomsky 1995b). In addition, phenomena exhibited by data such as (23) (Lasnik and Saito 1984, 1992) are also accounted for.

(23) Who do you think \([I_{\text{wh}} C_{\text{that}} [\text{John said } [I_{\text{wh}} C_{\text{null}} [I_{\text{wh}} \text{ left}]]]]\]

\(^7\) Epstein (1987, 1991) sought to deduce this asymmetry from an independently motivated A versus A-bar asymmetry in indexing, independently proposed in Chomsky’s (1982) analysis of parasitic gaps (which are licensed by A-bar, but not A binding). Epstein’s (1987, 1991) analysis nonetheless resorted to levels, binding, (asymmetric) indexing, \(\gamma\) -features, and stipulations regarding the phonological content of \(\text{that}\) versus \(C_{\text{null}}\).
In (23), a wh-element “crosses” two CP phase cycles. In the lowest CP phase cycle, the wh-element moves to SpecTP and SpecCP, so the deletion rule applies to SpecTP; hence, only $C_{null}$ is permitted in the embedded $C$. In the next CP phase cycle, however, the wh-element moves directly to SpecCP, not to SpecTP, so the deletion rule does not apply to SpecTP. SpecTP is phonologically realized as /John/, and the local $C$, namely $C_{that}$, is phonologically realized as /that/.

### 4.3. Cross Linguistic Variation of Complementizer-Trace Effects

As our deductive analysis now stands, there is a problem. As is well-known, complementizer-trace effects are parameterized, i.e. allowed in some grammars. Our analysis, based on deep principles such as Recoverability and Minimal Computation, presumably un-parameterized, would seem overly restrictive, wrongly predicting that that-trace effects are universally barred. Here we attempt to address this problem of how that-trace can possibly ever be allowed.

To begin, consider the well-known que-qui alternation. Roussou (2010) notes that it is possible to analyze qui as que + il (Rooryck 2000) or as que followed by the clitic -i (Taraldsen 2002). Given these possibilities, consider the following contrast (Roussou 2010:108):

(24) Qui penses-tu qui/*que est venu?
    who think-you that / is come
    “Who do you think has come?”

If a pronominal form (such as il or -i) counts as a phonologically realized SpecTP at some point of a derivation, then there is no reason for the deletion rule to apply to SpecTP, and $C+$SpecTP is phonologically realized as complex /qui/.

If SpecTP is literally absent, as argued in pro-drop languages e.g. Italian, then the deletion rule, by definition, cannot apply to SpecTP. If there is no deletion in SpecTP, there will be no corresponding deletion of a $C$ occupying the same position as SpecTP. Given this, consider (25) (Roussou 2010:106):

(25) Chi credi che abbia telefonato?
    who think-2s that has telephoned
    “Who do you think has telephoned?”

If $chi$ directly moves to (embedded) SpecCP (Rizzi 1982), i.e. not through SpecTP, then, in principle, nothing prevents $C$ from being phonologically realized as /che/. But the literal absence of SpecTP would circumvent the “two-peaked” analysis of phase-based application of Transfer. That is, under EKS (2012), “countercyclic” creation of SpecTP is precisely what triggers Transfer (to resolve the “two-peaked” situation). One possible way for us to guarantee cyclic Transfer is to pursue the line of Alexiadou and Anagnostopoulou’s (1998) analysis, where the inflection is understood to correspond to the morphological realization of EPP. Specifically, if there is (EPP satisfying) “countercyclic” verb movement to $T$ with rich agreement forms \{V, \{T, vP\}\}, then it will induce a “two-peaked” or intersecting situation, thereby triggering Transfer, as desired for all “finite clauses” (even those lacking SpecTP).

Finally, consider the following pair, which raises a paradoxical problem: just as in (26), in (27) the deletion rule does apply to SpecTP, but $C_{that}$ is nonetheless allowed, and it is phonologically realized as /that/ (Culicover 1992, Browning 1996):
Robin met the man that Leslie said that was the mayor of the city.

Robin met the man that Leslie said that [for all intents and purposes] was the mayor of the city.

In the derivation of (26), $C_{\text{that}}$ must undergo deletion (together with $\text{SpecTP}$), but $C_{\text{that}}$ bears phonological features and it is the sole copy. Thus, the required deletion of $C_{\text{that}}$ violates Recoverability. By contrast, in the derivation of (27), $C_{\text{that}}$ is curiously allowed, given the adverbial. The obvious difference between (26) and (27) is that an adverbial expression *for all intents and purposes* appears between $C_{\text{that}}$ and SpecTP, but how does this difference receive a principled account?

Browning (1996) argues that (27) differs from (26) in the structure of the embedded CP. She assigns to (27) the following CP recursion structure, where crucially there is complementizer movement from C-to-C:

$$[\text{CP} \quad \ldots \quad C \quad [\text{CP} \quad [\text{for all intents and purposes}] \quad [t_{\text{C}} \quad \ldots \quad ]]]$$

If Browning’s independently motivated CP recursion (or C-movement) analysis is correct, then the observed contrast, we argue, naturally follows. Consider the relevant aspects of the CP recursion structure of (27), illustrated in (29):

$$[\text{CP} \quad t_{\text{op}} \quad C_{\text{that}} \quad [\text{CP} \quad [\text{for all intents and purposes}] \quad [t_{\text{C}} \quad [\text{TP} \quad t_{\text{op}} \quad \text{was the mayor of the city}]\ldots ]]]$$

In the derivation of (27), the merged sisters of the lower occurrences of the operator OP are all deleted. There are (at least) two lower occurrences of OP in the embedded CP: \{T, \text{vP}\} and \{\text{v}, \text{VP}\}. The merged sisters of these occurrences include two copies of OP and one copy of $C_{\text{that}}$. Recall that, under the “two-peaked” or intersecting situation, the lower copy of $C_{\text{that}}$ counts as a merged sister of \{T, \text{vP}\}, and it undergoes deletion (together with SpecTP and SpecvP). Recall, in the core, ungrammatical *that*-trace effect (11), this results in an obligatory but unrecoverable deletion of *that*. But notice, if the lower copy of $C_{\text{that}}$ gets deleted in (29), the phonological features of $C_{\text{that}}$ are recoverable, because the higher copy of $C_{\text{that}}$ still remains. Thus, the deletion of the lower copy of $C_{\text{that}}$, forced by Minimal Computation, satisfies Recoverability, predicting that if *that* moves to a higher position, there will be no *that*-trace effect involving the departure site, exactly as in the derivation of (27), informally represented as (29).

5. **Summary**

We first reviewed simplest Merge and its generation of “doubly-rooted” or set-intersection structure, and then how Minimal Computation simplifies transfer to SM in accord with Recoverability. Recoverability requires $C_{\text{that}}$ to be phonologically realized as /that/, and Minimal Computation requires the lower copies of who to get deleted (since only one set of phonological features of who entered the derivation, so by Recoverability only one can appear at SM). These two (opposing) very general requirements, Minimal Computation (= “not too much”) and Recoverability (= “not too little”), however, appear to yield the *that*-trace effect applying only when the deletion rule applies to the merged sisters of \{T, \text{vP}\}, which include both (local) $C_{\text{that}}$ and SpecTP. As we noted at the outset, our analysis is unique in incorporating binary simplest Merge, yet it generates multiple sisters of a single category, as

---

8 We take deviance of (26) to mean that the deletion algorithm does not distinguish wh-elements such as who and a null operator OP.
discussed. A final oddity is that surprisingly, under the “two-peaked” analysis, C\(\text{that}\) and SpecTP are not in fact syntactically related, as can be seen clearly in (12) and (14)/(15), i.e. C does not c-command local SpecTP nor does local SpecTP c-command C. In current terms neither can minimally search the other, hence we have provided an account of complementizer and subject-trace effects, which recognizes no syntactic relation between the two positions. Hence, our account of that-trace makes no appeal to a syntactic relation from C to its local SpecTP. However, (a local) C and SpecTP each counts as a merged sister of \{T, vP\}, under the “two-peaked” analysis, employing simplest Merge (3i) barring both replacement (Freidin 1999) and label projection (Chomsky 2012). Given this, we argued that the core that-trace contrast, exhibited by (10) and (11), was deducible from the two very general principles Minimal Computation and Recoverability, neither of which appears to be specific to a particular construction or language, and may be in part or in full, not-specifically-linguistic principles with far more general application in computational (and/or biological, and/or physical) systems of other types that (informally) bar “random deletion and insertion.” That is, perhaps these are both good candidates for third factor laws, hence linguistic deduction from more general non-linguistic principles. But that important, fascinating and unavoidable property of all scientific inquiry – concerning the level at which explanation occurs (see Chomsky (1965:51, 2005) for bringing this aspect of science to Linguistics) – awaits further, interdisciplinary collaboration, given our ignorance concerning the broader issues at hand. We then gave brief preliminary analyses of a cross-linguistically variant range of grammatical complementizer-trace phenomena including the que-gui alternation (Roussou 2010, Rooryck 2000, Taraldsen 2001), pro-drop languages like Italian (Rizzi 1982), adverb that-trace phenomena (Culicover 1992, Browning 1996), and so-called ECP asymmetries (Huang 1982, Lasnik and Saito 1984, 1992, Safir 1985, Chomsky 1991, Rizzi and Shlonsky 2007).

References

Interfaces + Recursion = Language?, Mouton de Gruyter, Berlin, 1-29.


1. Preliminaries: Unselected embedded questions

This paper aims to clarify under which conditions questions can be embedded. It concerns the syntactic and semantic selection (so-called c- and s-selection) of verbs of propositional attitude, i.e. verbs denoting mental process towards a class of abstract objects that are propositions. A question is related to an interrogative speech act. It is therefore embeddable under an interrogative performative verb as shown with *ask* in (2), the embedding of (1). Questions are also embedded under interrogative non performative verbs as in (3) with *wonder*. More surprisingly, they also appear with non-interrogative verbs such as *know* in (4)a (though not in all contexts, see (4)b).

(1) Did Peter come?
(2) I ask you if Peter came.
(3) I wonder if Peter came.
(4) a. I don’t know if Peter came.
   b. ?I know if Peter came.

*Know*, cognitive verbs and declarative verbs are not interrogative verbs, they do not denote the search for an answer. With them, the embedded clause denotes rather the answer to the question than the question itself. Note that the DP *the answer* can substitute for the *if*-clause in (4). This dichotomy between question embedders has been noticed for a long time. Non-interrogative question embedders are dubbed responsive predicates by Lahiri 2002 and resolutive predicates by Ginzburg 1995, a class that corresponds to Groenendijk and Stokhof’s 1984 extensional predicates. Resolutive predicates are predicates that otherwise select for *that*-clauses as shown in (5).

(5) I know that Peter came.

According to most theories, a *that*-clause denotes a proposition. *Know* then semantically selects for propositions. On the other hand, a question is taken to denote in the most influential theories either set of propositions <<s, t>, t> (Hamblin 1973, Karttunen 1977) or a proposition meaning <s, <s, t>> (Groenendijk and Stokhof 1984), but not a proposition (type <s, t>). This means that we are dealing either with two homonym verbs *know* or with an s-
selection mismatch in (4)a. The former option is not desirable, for *know* in (4)a and (5) intuitively denotes the same mental process. The path of the s-selection mismatch is more appealing. It is the main topic of two articles, Adger and Quer 2001 and Oehl 2007 and is now known as the unselected embedded question (UEQ) issue.

In the present paper, I shall argue that both accounts wrongly propose that the mismatch is solved by an external operation (interrogative raising for Adger and Quer, reconstruction of the main predicate for Oehl), rendering the main clause suitable for *yes/no* question embedding. They are also weakened by the fact that they are not extendable to *Wh*-questions, which should be precluded after the same set of verbs but, unexpectedly, are not. In their view (6)a should not be fine since *knew* is not in a nonveridical context (see below).

(6) a. Peter knew who came at the party yesterday.
   b. Peter doesn’t know who came at the party yesterday.

To account for that, they need further machinery. I shall here concentrate on the first issue (*yes/no* questions) and show that the operation rather takes place in the embedded clause through an operator, syntactically detectable. I shall briefly suggest at the end that it extends to the *Wh*-question issue, leaving the complete proof for future research.

The article is organized as follows: in section 2, I review the conditions under which UEQs are licensed; section 3 highlights the limitations of the previous studies on the issue; in section 4, I take it up anew from the syntactic side, showing that an operator surmounts UEQ-clauses; section 5 explains what the semantic role of this operator is. Section 6 draws some conclusions from this study and suggests some application to *Wh*-UEQs.

2. *Yes/no* question licensing in nonveridical contexts

   Both Adger and Quer’s 2001 and Oehl’s 2007 theories rest on the same observation. In English, we notice the contrast in (7)a/b (inspired by Adger and Quer 2001).

(7) a. Peter admitted *if/ok*that Julie was sober.
   b. Did Peter admit *ok*if Julie was sober?

   A *yes/no* embedded question is barely acceptable after *admit* in a positive declarative sentence (7)a, but it is natural when the sentence is a question (7)b. Note that in (4)a, *know* is negated. All this suggests that special conditions are in order for resolutive verbs to embed questions, contrary to interrogative verbs: both (8)a/b with *ask* are perfectly acceptable.

(8) a. Peter asked you if Julie was sober.
   b. Did Peter ask you if Julie was sober?

   The aim of those papers was to clarify the conditions under which resolutive verbs embed questions. They notice that, along with negation (4)a and interrogation (7)b, other contexts improve more or less the acceptability of the sentences (depending on the predicate), like antecedents of conditionals (“*If I knew if* Peter was coming, I would have been happy”). Adger and Quer point out that these conditions are also conditions for NPI licensing. In (9) through (12) (adapted from Adger and Quer 2001:111), the NPI *ever* is licensed under the scope of negation and interrogation ((9) and (10)), but not in positive contexts or outside the scope of the negation ((11) and (12)).

(9) You didn’t **ever** go to London.
(10) Did you **ever** go to London?
(11) *You ever went to London.

(12) *You ever didn’t go to London.

The list of operators licensing NPIs extends to negative quantifiers, only focus, antecedents of conditionals, adversative predicates, without-clauses for English, and possibly to other environments in other languages, depending on the language sensitivity. There is a long-standing debate on what feature(s) these environments have in common: are they downward-entailing or nonveridical? The first hypothesis faces issues although it has been rescued by von Fintel 1999. But it is not comprehensive enough for all languages as shown by Giannakidou 1998. After a survey of Modern Greek data, she notices that some non-downward-entailing environments such as future tense or imperative license NPIs in this language. That is why she proposes to borrow Zwarts’ 1995 notion of nonveridicality to account for NPI licensing. (13) is her formulation of Zwarts’ definition of nonveridicality.

(13) (Non)veridicality for propositional operators
   i. A propositional operator $F$ is veridical iff $Fp \rightarrow p$; otherwise $F$ is nonveridical.
   ii. A nonveridical operator $F$ is antiveridical iff $Fp \rightarrow \neg p$.

This leaves us with the following list for Modern Greek: Negation; negative quantifiers; questions; downward-entailing operators; future tense; imperatives; habitual/generic sentences; (certain) modal verbs; adversative predicates; directive intensional verbs; protases of conditionals; without-clauses; before-clauses; comparative clauses; only focus; disjunction.

This list mostly overlaps with the list of NPIs licensers in English. From now on I will take veridicality to be the condition for yes/no question embedding under resolutive predicates in English1. Adger and Quer 2001 and Oehl 2007 show that this is also true of Basque, Catalan, Spanish, German and Persian. This also holds of French ((14) translates (4) and (5)). (15), though somewhat literary, shows that at least negation and interrogation are NPI licensers in French.

(14) a. Je ne sais pas s’il est venu.
    b. ?Je sais s’il est venu.
    c. Je sais qu’il est venu.

(15) a. Pierre n’a jamais été à Londres.  b. Pierre est-il jamais allé à Londres?
    ‘Pierre has never been to London.’     ‘Has Pierre ever been to London?’

We now know more precisely that resolutive predicates need association with a nonveridical operator to be able to embed questions. This said, we still are not able to explain how the selection mismatch is possible between the embedder and the embeddee. In the next section, we review three proposals concerning this issue.

---

1 Be that as it may, the theory of NPI licensing adopted here does not matter. It is sufficient for the proof that NPIs and UEQs have the same distribution.
3. Previous analyses

3.1. Adger and Quer’s 2001 proposal

As said in introduction, Adger and Quer 2001 notice that questions are not expected with resolutive predicates and attempt to provide an explanation for this fact. First of all, they point out that in this very case, in some languages, a morpheme shows up in addition to the interrogative complementizer. For example in Basque, the interrogative complementizer in yes/no embedded question is -(e)n ((16) = their (45)²). But with resolutive predicates in nonveridical contexts, it is -(e)n + ik ((17) = their (47)) is an example of negation, but they also provide examples with question and antecedent of conditional).

(16) [telebistako langileek grebe egingo dute-n]
television-Gen workers-Erg strike-abs make-Fut Aux-3Pl.Erg.3Sg.Abs=Comp
galdetu
diet.
ask-Ptcp Aux-1Sg.Erg.3Sg.Abs.3Pl.Dat
‘I have asked them whether the TV workers will go on strike.’

(17) Jonek ez du esan
Jon-Erg Neg Aux-3Sg.Erg.3Sg.Abs say-Pret
[Bilbora joango d-en-ik]
Bilbao.all go-Fut Aux-3Sg.Abs=Comp=ik
‘Jon didn’t say if he was going to Bilbao.’

They then focus on the morpheme -ik. They analyze it as a specific polar determiner that heads the embedded clause and that is sensitive to the environment provided by the matrix. It enters into a licensing relation with nonveridical operators. That is why, when there is no nonveridical operator in the environment, the clause is not acceptable.

The second step of their proposal is that this explanation carries over to English. In English as well, one can find such polar determiner, but it is covert. Nevertheless, its presence with a matrix resolutive verb in a nonveridical context is shown by an extraction test. It is not possible to extract a Wh-phrase out of a DP, especially out of a DP headed by an NPI D, such as (18) (the examples are theirs). This is not true of if-clauses after an interrogative verb such as ask in (19)a, but it is true of (19)b after a resolutive verb such as admit.

(18) *Who didn’t you believe [any rumor that John killed tij]?

(19) a. ? What did no one ask [if John had stolen tij]?
    b. *What did no one admit [if John had stolen tij]?

This suggests that if-clauses with admit behave as DPs headed by an NPI, i.e. like their equivalent in Basque. Actually, the judgments are rather fuzzy, though most French informants I asked feel the same (slight) contrast between (20)a and (20)b (the French version of (19)a/b).

(20) a. ?? Qu’est-ce que personne n’a demandé [si Jean avait volé tij]?
    b. *Qu’est-ce que personne n’a reconnu [si Jean avait volé tij]?

² I keep the glosses of borrowed examples as in the original version, although I clarify some abbreviations by replacing them with the Leipzig glossing rules.
But there is a worse objection. *If* in English and -(e)n in Basque already have an NPI distribution. They appear in nonveridical contexts and with interrogative verbs (i.e. a specific case of nonveridical contexts, for they license NPIs as well as direct questions do, see (21)).

(21) I asked him if he had ever been to Monaco.

What we expect would be something to repair the mismatch between an attitude verb and a question and not between the nonveridical operator and the question. Put otherwise, the features of the nonveridical operator and those of the complementizer already match, another NPI is needless. The use of -ik can be explained if we take into account the features that Basque Cs spell out, as I do in section 4.3.

Another problem is that Adger and Quer account for the issue in terms of licensing, while we are dealing with a c- and s-selection issue. This is the gist of Oehl’s criticism, albeit Adger and Quer propose to resolve the s-selection mismatch by the raising of the interrogative to the restrictor of a quantifier (see section 5.1). All these criticisms show that something must be improved in Adger and Quer’s theory. I shall leave it aside for the moment and come back to their data later.

3.2. Oehl 2007

Oehl 2007 points out that it is unexpected that the s-selection of a clause depends on the context and not on the sole meaning of the embedding predicate. If resolutive predicates embed propositions, there is still a mismatch to explain between the embedder and the embedded clauses. Moreover, Adger and Quer’s proposal necessitates that some embedded clauses are CPs and other are DPs, a non-desirable result according to him. He proposes that the phenomenon of UEQ is due to a change in the meaning of the embedding predicates, that become *if*-embedders, and not to the structure of the embedded clause. This change in meaning is made possible by several factors.

First, these predicates are cognitive factive predicates (admit being a borderline exception). With such predicates, the presupposition of the complement is weaker than with emotive factive predicates (Karttunen 1971, Faure 2006). Put otherwise, under certain (nonveridical) circumstances, the matrix (cognitive factive) predicate does not trigger the presupposition of its complement and the embedding of *if*-clauses becomes possible.

Second, he proposes that *if*-clauses are not interrogative complementizers, but nonveridical operators (Mod°) that indicate that the proposition denoted by the clause does not have a truth value. If there is a nonveridical operator in the matrix clause, these operators can be employed. The same operators are used in clauses denoting indirect questions. This means that they are underspecified. In other terms, there is no need for a polar D since the C is

---

3 Except in one context for Basque -(e)n, see below section 4.3.  
4 This approximation is sufficient for the purpose of the present paper, but it is actually not true. The relevant notion for question-embedding by resolutive predicates is the veridicality of the predicate and not its factivity, as shown by Égré 2008 on English and French data. Working on French and Classical Greek data, Faure 2010 adds that for a question to be embedded under a veridical predicate, this predicate must also have an assertive feature. Otherwise, I fail to see why emotive factive predicates, which are also veridical (pace Égré), cannot embed questions across languages.  
5 Actually, we still need the factivity to be active since resolutive predicates embed clauses that denote the TRUE answer to the question and not merely a possible answer.  
6 In this view, it is not surprising that in many languages (in English, French, Greek etc., but neither in German nor in Basque) these operators are also used as complementizers in antecedent of conditionals.
itself polar (recall the criticism of Adger and Quer’s position in the previous section).

The gist of Oehl’s proposal is his attempt to rescue the notion of s-selection. According to him, the nonveridical operator in the matrix clause attaches to the matrix predicate to form a complex predicate that selects the nonveridical embedded clauses, and not, crucially, that licenses them. On the basis of German data, he draws a parallel with restructuring verbs or internally negative verbs. He appeals to a parallel that changes the s-selection. For example in (22), two options are a priori available. Either the ob-clause has been licensed by the negation nicht or it has been selected by the complex predicate nicht-gezeigt. This makes different predictions regarding topicalization. In the first case, the verb can move up without carrying along the negation (b), in the second case, it must carry it up to the left-periphery (c). The latter option is more acceptable than the first one, suggesting that the negation is attached to the verb and that they form a whole, with a specific s-selection.

(22) a. Er hat eigentlich nicht gezeigt, ob das stimmt.
   he has actually not shown if this is-true
   shown if this is-true has he actually not
   not shown if this is-true has he instead actually

This proposal bridges the selection gap. Unfortunately, judgments are even more fuzzy about the contrast (22)b/c than about (19)a/b. My informants report that it is more noticeable in Southern Germany than in the North. On the other hand, it does not account for Adger and Quer’s 2001 extraction contrast provided in (19)a/b.

In addition to this, note that for some languages, such as Classical Greek (CG), this proposal is not applicable. The configuration of the clause does not allow for such a restructuring. In sentences like (23)7, when the negation and the predicate would undergo restructuring, the predicate is no longer adjacent to the UEQ. The new predicate and the UEQ do not form a constituent, precluding a German-like topicalization. Worse, at the step where the UEQ is merged, it is not yet properly selected, because it merges with the predicate before the predicate merges with the negation, i.e. before the restructuring of the predicate that should select it. Oehl’s proposal thus leads to important difficulties in the derivation of the clause.

(23) Hósai mè faneròs èn hòpòs egígnòskeni, …(X. Mem. 1.1.17)
   ‘lit. (They did not know) what he (Socrates) is not obvious how he had decided, …’
   ‘(They did not know) opinions of his that were unknown to them.’

This means that, though interesting, this proposal has a narrow explanatory power. That is why we must discard the restructuring idea, while keeping the remark on the complementizer’s polarity.

---

7 Note that this example could also be an objection to Adger and Quer’s proposal. Their theory forbids extraction out of a clause embedded under a resolutive verb. Still, the Wh-term hósai has been extracted out of the hòpòs-clause. Nevertheless, the properties of extraction are poorly understood in CG and this language is very permissive in this respect, allowing structures precluded by the theory like ‘he is clear that t is nice’ (but not allowing extraction out of a complex DP).
3.3. Roussou’s proposal

Roussou 2010 agrees with Oehl that there is no need to postulate an NPI D above the if-clause since the distribution indicates us that if itself plays this role. Her proposal is that if and that are in fact Ds that select for CPs. Thus all completive embedded clauses are in fact DPs and there is no syntactic mismatch. The issue therefore boils down to the semantic mismatch. This latter mismatch is solved by postulating that if returns a proposition (without a truth value) and not a question. In this case, there is no semantic mismatch between, say, know which requires a proposition and the if-clause8. The if-clause is a kind of NPI that needs licensing by an interrogative verb (in the case of ask) or by a nonveridical operator (in the case of resolutive predicates).

However, this hypothesis faces for some empirical objections. In Roussou’s view, the DP headed by if is then directly selected by the matrix verb. This predicts that there is no intervener between the embedded clause and the matrix verb. But many languages present a second complementizer above the equivalent of if, like Persian or Spanish, discussed in section 4.2 (see already Basque and English according to Adger and Quer 2001).

3.4. Summary

In this section, we challenged three proposals to account for the mismatch issue that UEQs represent:

— [Adger and Quer 2001] UEQs are surmounted by a polar D licensed by a nonveridical operator and the s-mismatch is solved through raising of the interrogative to the restrictor of a quantifier. This proposal does not take into account the nonveridical feature of the UEQ C; it is maybe not suitable for CG (see footnote 8).

— [Oehl 2007] The nonveridical operator and the main verb undergo restructuring and select for an interrogative. This proposal is not suitable for Northern German in the first place and involves flaws in the derivation of the clause.

— [Roussou 2010] The complementizer if is actually a polar D that is licensed by the nonveridical operator and if-clauses denote propositions. This proposal is not suitable for Persian, Basque or Spanish that present functional heads above the interrogative C.

They all face a common issue. As mentioned in introduction, embedded Wh-questions and yes/no questions should have a uniform treatment, which is not taken care of in these approaches9. Moreover, each proposal applies only to certain languages. Given that the UEQ issue is first of all a semantic problem, it should be able to receive an answer generally applicable across languages. As we shall see in the next two sections, the pattern is common to both the syntax and the semantics of the languages under discussion.

4. A multi-layered complementizer system

Adger and Quer 2001 posited a D above the interrogative C. The split of the left periphery of the clause is reminiscent of the so-called cartographic hypothesis. In this hypothesis, the CP domain hosts several functional projections relative to the pragmatic functions, such as illocutionary force, topic, focus etc. In this section we try to flesh out the division of labor between the different projections of the left-periphery of the clause and what

---

8 Note nevertheless that this idea means that ask also selects for a proposition, and not for a question, which is not an uncontroversial claim.

9 Oehl suggests that Wh-questions are licensed under conditions different from those of yes/no questions, namely when in the scope of focus. This suggestion is at odds with the data, for a Wh-question can be topicalized, thus scoping over a focus operator.
Adger and Quer’s D could be identified with in order to better understand what its role is.

4.1. The interrogative C in the cartographic hypothesis

In a 2001 paper, L. Rizzi addresses the issue of the interrogative C and Wh-item positions. He attempts to incorporate them into the structure he proposed in his 1997 seminal article. (24) through (26) are the crucial data. Whereas che ‘that’ cannot be surmounted by a topic DP, se ‘if’ can, which means that se is located in a projection below che.

(24) *Maria crede, il tuo libro, che lo potrà leggere.
  ‘Maria believes your book, that (she) will be able to read it.’

(25) Maria crede che, il tuo libro, lo potrà leggere.
  ‘Maria believes that, your book, she will be able to read it.’

(26) Non so, a Gianni, se avrebbero potuto dirgli la verità.
  ‘I don’t know, to Gianni, if they could have said the truth.’

Rizzi dubs the complementizer che Force° and se Int°. He places them relatively to the other projections:

(27) Rizzi’s 2001 left-periphery of the (embedded) clause10, 11

\[\text{FORCE(C1) (TOP*) INT(C2) (TOP*) FOC Wh (TOP*) FIN IP}\]

As ‘interrogative’ is also an (illocutionary) force, these labels do not sound felicitous to me. For the moment, let us adopt a neutral position by labeling FORCE “C1°” and INT “C2°”.

4.2. Languages with overt C1 complementizers

Both in Adger and Quer’s and in Rizzi’s theory, there is a lower C endowed with the feature /interrogative/. In each theory it is surmounted by another position, filled or not by a functional head: an NPI D for Adger and Quer, Force for Rizzi. A welcome result would be to show that the two functional heads are one and the same since they share the property to be active with a declarative verb and non active with an interrogative verbs (recall the analysis of (19)a/b). Despite these common features, the two proposals are hardly compatible. For Rizzi, C1° is declarative, for Adger and Quer, it is an NPI D. It seems that we need to rework the structure of the left-periphery of the clause to incorporate Adger and Quer’s NPI D.

If we look at other languages, some present several overt complementizers. This is the case for Basque and Spanish (see examples (17) and (30), borrowed by Rizzi 2001 from Margarita Suñer), but also for Persian and Dutch (examples (28) borrowed from Oehl’s 2007 (42)b and footnote 19; (29) borrowed from van Craenenbroeck 2010:30).

(28) a. (man) nemīdainam ke āyā ī zabānšenāsī mīxānad.
  I Negknow Subordination Q he/she linguistics studies

---

10 The stars indicate recursive projections.
11 Rizzi’s data regarding the relative position of Wh-items, contrastively focused material and the interrogative C are disputable and there are clues that these positions might be merged into only one (see López 2009), but this is irrelevant to our topic.
‘I do not know if he/she studies linguistics.’

b. man dīdam ke ān pesar zabānsenāsī mīxānad.
I saw Subordination Dem boy linguistics studies
‘I saw that the boy studied linguistics.’ [Persian]

(29) Ik weet niet of dat Jan gaat komen.
I know Neg if that John goes come
‘I don’t know if John will come.’ [Dutch]

Nevertheless, not all these functional heads are candidates that can be identified with Adger and Quer’s 2001 NPI D, mostly because they have a different distribution.
In Spanish, the co-occurrence of *que* with an interrogative C hinges on very strict factors as shown by Lahiri 2002. It is not a declarative marker, but a quotation marker, available with and only with speech act verbs. Thus, a cognitive verb like *saber* ‘know’ or a verb of inquiry like *investigar*, which are not speech act verbs, do not allow for *que* to precede *si*, as in (31) and (32) (adapted from Lahiri 2002:chapter 6, examples (51)a and (64)d).

(30) Me preguntaron (*que) si tus amigos ya te visitaron en Granada
me they.asked that if your friends already you visited in Granada
‘They asked me that if your friends had already visited you in Granada.’

(31) El detective no sabe (*que) si Pedro la mató.
the detective Neg knows that if P her=killed
‘The detective doesn’t know if Pedro killed her.’

(32) Investigarán (*que) si se puede curar el SIDA.
they.investigate that if Refl=can cure-Inf the AIDS
‘They will investigate if AIDS can be cured.’

(33) Dijo que a no molestarle.
he.says that P Neg bother-Inf=him
‘He said not to bother him.’

The quotation marker *que* is not involved in the insertion of UEQ and is possibly not to be identified with the assertive complementizer *que*. The latter is limited to finite clauses, whereas the former can show up in nonfinite clauses, provided that they report a speech act, as can be seen from (33) (Lahiri 2002:chapter 6, example (68)a), where the infinitive subordinate clause report a speech act of order. Nonetheless, the quotation marker and the assertive C do not show up together as expected, maybe for some reason precluding the sequence *que*1 *que*2. If we are on the right track, this means that C1P hides in fact several functional heads: a quotation marker and an assertive C. We end up with the following structure:

(34) Structure of the left-periphery of the clause (1) (to be revised)
[C1°(quot) … C°(assert) … C2°(int) … IP]

Persian presents us with a different pattern. Contrary to Spanish *que*, the Persian complementizer *ke* is not limited to a subset of question-embedding verbs but shows up with

12 Note that it is grammatical when *que* is an *Wh*-item (written *qué*).
all verbs, including interrogative ones (35) (Oehl’s (42)a).

(35) ū porsīd ke āyā man zabānšenāsī xānde būdam.
    he/she asked Subordination Q I linguistics studied had
    ‘He/she asked if I had studied linguistics.’

As seen in (28)b, ke is also the form that the assertive C takes. A priori, the two ke should not be identified with one another, for they have a different distribution (the assertive ke cannot show up without āyā after interrogative predicates) and then serve different functions. In (28)a and (35), ke1 is a subordination marker since available whenever finite subordination is at stake, in (28)b, ke2 is an assertive C. This means that there could be a position for the subordination marker in the CP domain. As this subordination marker is limited to finite clauses, it could be located in Rizzi’s Finite projection, at the bottom of the CP domain. We thus end up with the following structure:

(36) Structure of the left-periphery of the clause (2) (final)

\[ C1°(quot) \ldots C°(assert) \ldots C2°(int) \ldots C''°(sub) \ldots Mod°(polarity) \ldots IP \]

If this proposal is on the right track, āyā would not be in C2°, but below, in a modal polar projection (as argued for by Oehl 2007). In some languages this modal polar particle could be located higher in a complementizer position, namely C2°, devoted to the interrogation. This is arguably what happens in Dutch.

Dutch is at first sight even more puzzling. Like Persian ke, the complementizer dat1 is used with all verbs in embedded questions (including ‘ask’, see (37), adapted from van Craenenbroeck 2010:30) and is homonymous with the assertive C dat2.

(37) Ik wraag me af of dat Jan gaat komen.
    I ask me PRT if that John goes come
    ‘I wonder if John will come.’

On a par with Persian ke, dat1 is therefore a subordination marker and not a quotation marker. As shown by van Craenenbroeck (2010:Chapter 4), its role is minimal. Its specifier is the place where operator features and variable-dependencies are checked, which basically means that it establishes a relation between the interrogative CP projection above it and the rest of the clause. What is noteworthy is that dat1 comes after the interrogative C. However this is not surprising provided that we assume the results achieved on Persian data. Dat1 is located in the C''° position and of is in C2° since it does not only mark the sentence as polar (as does Persian āyā), but as interrogative.

In this section, we saw that several projections surround the interrogative C. All these heads can be spelt out as a C. If we superpose Rizzi’s structure with ours, we end up with the following result:

\[
\begin{align*}
    & [C1°(quot) \ C°(assert) \ C2°(int) \ C''°(sub) \ Mod°(pol) \ IP] \\
    & [\text{FORCE(C1)} \ (\text{TOP}*) \ \text{INT(C2)} \ \text{FIN} \ IP]
\end{align*}
\]

This is beyond the aim of this paper, but we can tentatively suggest that Top° and Foc° can be spelt out as complementizers, as claimed in Ledgeway 2005 or Mascarenhas 2007. In this view, the diversity of complementizers would merely reflect the diversity of topics (frame setting topic, frame topic contrastive topic, continuous topic, see Lambrecht 1994), a frame setting topic being for example associated with the quotation or the force C.

Languages vary with respect to which of these projections they grammaticalize. In Classical Armenian, only one item spells out the different complementizers (including the

Regarding the UEQ issue, C** (quot) and C° (sub) cannot play any role, for their distribution is tangential to that of UEQs. Only C° (assert) remains as a possible candidate with which Adger and Quer’s D could be identified. Basque provides us with arguments for this identification.

4.3. Basque

In the relevant contexts of Basque, we deal with four items: -(e)la, -(e)n, -ik, -a (see (16), (17), (38)).

-(e)la is a declarative complementizer, apparently restricted to speech act declarative verbs (38) (= Adger and Quer’s (50)). It can be used when the proposition is pragmatically presupposed (the presupposition is not triggered but the proposition already belongs to the common ground).

(38) Jonek ez du esan [Bilbora joango d-ela].
Jon.Erg Neg Aux.3Sg.Erg.3Sg.Abs said Bilbao.ALL go.Fut Aux.3Sg.Abs=Comp
‘Jon didn’t say that he was going to Bilbao (although he is).’

More interesting is the complementizer -(e)n. It is available with interrogative (16) and resolutive (17) verbs (factive and veridical), but it can appear by itself only with interrogative verbs. With resolutive verbs, it shows up both with -ik (elsewhere used as an NPI D) or -a (elsewhere used as a definite D). With -a, it is equivalent to a presupposed propositions (in a triggering environment, contrary to -(e)la). With -ik, it is equivalent to a UEQ if-clause or to a that-clause that denotes a non-presupposed proposition. Apparently the role of -ik is vacuous since the meaning of the -(e)n-clause is the same as when it is absent. The only difference resides in the main verbs, ‘know’ selecting a proposition and ‘ask’ a question. This means that the sole role of -ik can be to mark the clause as subordinate to an assertive verb and not to an interrogative verb, while -a is responsible for both the clause typing and the truth value of the proposition denoted by the embedded clause. In this case, -ik and -a must be the equivalent of one of the complementizers of the structure in (36). Put otherwise, -ik seems to be the form that -a takes when it does not act on the -(e)n-clause. This resembles an agreement between the determiner and the complementizer, the determiner endorsing the nonveridical feature of the complementizer. According to their distribution, they cannot be C° (quot), for they appear with non-speech act verbs. They are not subordination markers, since they do not appear in every (finite) subordinate clause. But they are available with assertive predicates, which makes them good candidates for the assertive complementizer. This is in line with the semantic role they play as we shall see.

If we are on the right track, the functional head that Adger and Quer’s tests detected in English (see section 3.1) is also an instance of the assertive complementizer.

4.4. Summary

In this section, we explored more deeply the syntactic side of the UEQs. Following the cartographic hypothesis, we found out that the functional heads of the CP domain can be realized as complementizers or as determiners, depending on what the language grammaticalizes. We singled out a functional head that we dubbed assertive complementizer which is present when UEQs are involved. Thus, we followed Adger and Quer 2001 in postulating a functional head that makes UEQs possible, but recall from section 3 that we differ from them in attributing the necessity of a nonveridical environment to the nature of the interrogative head and not to this functional head (thus following Oehl 2007 and Roussou
5. A type-shifting operator

In this section, I shall claim that \( C^\circ \) is an operator that takes a question (whatever a question is) and returns a proposition, thus satisfying the selection properties of resolutive predicates. To defend this claim, I shall argue against Lahiri’s 2002 and Adger and Quer’s 2001 quantifier raising treatment of interrogatives.

5.1. Against an interrogative raising approach

Going back to the selection issue, there is a mismatch between the selection properties of resolutive predicates (which are, as said in introduction, proposition-selectors, i.e. are of type \(<s, t> <e, t>>\) and questions that are of type \(<s, t> t>\) (in the most influential theories, Hamblin 1973, Karttunen 1977\(^{13}\)). To solve this issue, Lahiri 2002 resorts to a process of Interrogative Raising (IR) triggered by Quantifier Raising (QR). This process applies when there is no way to solve the type mismatch such as in UEQs. Let us look at (39) (Lahiri’s (31)a, p. 77).

(39) John knows, for the most part, which students came to the party.

(40) (a)[[CP which students came to the party], (b)[most (c)[IP John knows ti]]].

In this operation, the sentence containing a UEQ has a quantifier, possibly overt, like \textit{most} in the example. Quantifiers are in this theory polymorphic. When they quantify over questions, they are of type \(<<<s, t>, t>, <<s, t>, t>, t>>\). The CP denoting a question rises at LF to the specifier of this quantifier adjoined to the IP, thus forming the LF in (40). The IP is in the nuclear scope of the quantifier and the CP(interrogative) in its restrictor. Crucially, the raised CP leaves a trace that is interpreted as the type required for functional application, i.e. \(<s, t>, t>,\) by a type-adjusting rule that Lahiri borrows from Bittner 1994 (see p. 83). As the node (b) of the LF contains a free variable of type \(<s, t>,\) the node undergoes a \(\lambda\)-abstraction and is of type \(<<<s, t>, t>, t>>\). By functional application, the quantifier applies first to the IP yielding a node of type \(<<<s, t>, t>, t>, then to the question, yielding something of type \(t\) as expected.

Lahiri grounds his hypothesis in the distribution of the Quantificational Variability Effect (QVE). The QVE was first described in Berman 1991. He points out that in sentences such as (39), the main clause quantifier \textit{most} unexpectedly binds the \textit{Wh}-variable, such that (39) means that the knowledge of John applies only to a subset of the answer propositions. The QVE, notices Lahiri, takes place only with resolutive verbs so that in a sentence displaying an interrogative verb like (41) (Lahiri’s (50)a, p. 88), if it has a meaning at all, \textit{most} cannot quantify over the embedded clause.

(41) John mostly wonders which students came to the party.

This is in line with the mechanism (IR) he posits for the interpretation of UEQs with resolutive predicates: The QVE applies to \textit{Wh}-clauses just when an additional process is needed for the interpretation to be possible. That is why a quantifier is always required, even

\(^{13}\) Note that the mismatch also arises in a Groenendijk and Stokhof 1984 model, where questions are of type \(<<s, t>, t>>\). However, Adger and Quer’s and Lahiri’s theory of interrogative raising would fail to account of the raising of a question of such type, as acknowledged in Adger and Quer (2001:footnote 17).
if not overt. In this very case, it is universal by default (Lahiri 2002:91).

Adger and Quer 2001 resorts independently to the same process for yes/no UEQs. The NPI D they argue for is the quantifier that triggers the Interrogative Raising. Instead of Bittner’s type-adjusting rule, they appeal to Heim and Kratzer’s 1998 claim that the trace of a raised generalized quantifier has a simple atomic type, i.e. t (<s, t> if intensionalized) in our case. Another apparently unimportant difference is the level to which the quantifier adjoins. Indeed, according to Adger and Quer, the D patterns with the quantifier either since a yes/no question involves a choice between two items. Like either, it scopes below the negation, i.e. over the sole VP. More generally, given that nonveridical operators are VP modifiers, to be in their scope, the D must adjoin to VP and not to IP in Adger and Quer’s view. Ultimately, the selection problem is then solved in the same fashion as Lahiri’s.

However, despite their kinship, there are important differences between the two approaches. First of all, Lahiri deals with Wh-questions while Adger and Quer address yes/no questions. Second, Lahiri does not take into account the veridical nature of the environments where the UEQ phenomenon takes place. Let us then assume that it should be available in veridical as well as in nonveridical contexts. This assumption turns out false, as shown by (42) and (43), two cases of nonveridical environments.

(42) #Does John know, for the most part, which students came to the party?
(43) #John does not know, for the most part, which students came to the party.

The contrast between (39), on one hand, and (42) and (43) on the other hand, shows that Lahiri’s process of IR is only available in veridical contexts. This is a major discrepancy with Adger and Quer’s view according to which yes/no UEQs show up and IR happens only in nonveridical contexts. The two proposals are then not compatible. To accommodate their proposal to Wh-questions, Adger and Quer would have to assume a D heading Wh-UEQs. But once again, that would not explain why Wh-UEQs contrary to yes/no UEQs are also available in veridical contexts (recall the data under (6)). We leave this last question for future research (see however the conclusion) and focus from now on on yes/no UEQs.

Adger and Quer’s proposal to solve the s-selection mismatch is moreover hardly compatible with the results we reached in sections 3. We saw that the semantic properties attributed to the D by Adger and Quer (nonveridicality) were in fact those of the interrogative C, with which the D agrees. The D is rather an assertive operator, sometimes realized as an assertive C. In this case, there is no reason why it should not rise past the VP and adjoin to the IP. But the UEQ would then move past the nonveridical operator to reach the quantifier restrictor and then scope over its licenser. There is another technical objection to Adger and Quer’s proposal. They advocate the configuration [VP [Q Restrictor] Nuclear Scope]]. In this kind of situation, as pointed out by Lahiri (2002:84-85), the quantifier does not c-command its nuclear scope, probably resulting in a scope problem.

However alluring the Interrogative Raising proposal might be, it is not to be retained since it does not provide us with a uniform treatment of UEQs and faces technical problems. The role played by the functional operator heading the UEQ must be explained otherwise in a frame where the UEQ remains in situ.

5.2. UEQs are interpreted in situ

The previous section has cast doubt on the movement of UEQs. This suggests that they are interpreted in situ. In this case they would have to be of type <s, t>\(^{14}\). In section 4 I

\(^{14}\) For the necessity to posit several types to account for questions see Groenendijk and Stokhof 1989.
proposed to identify Adger and Quer’s D with an assertive functional head. This head is present with resolutive predicates. Recall that resolutive predicates are factive and veridical predicates\(^{15}\), which means that the proposition denoted by the embedded clause is presupposed or is a true proposition\(^{16}\) (or a fact in Ginzburg’s 1995 view). For this reason, I argue that the role of the functional head present with such predicates is to turn the embedded question into a true proposition that is the true answer to the question, thus resolving the type mismatch.

This proposal was already made by Chierchia (1993:192-194), but in different circumstances. He discusses the QVE described in the previous section. He adheres to Lahiri’s proposal (which dates back to 1991) of IR, except in the case described after example (41). According to him, in absence of an overt quantifier, there is no covert universal quantifier. In this case, the type-mismatch between the UEQ and the resolutive verb remains. Chierchia proposes to solve it by a type-shifting operated by the “supremum operator” \(\sigma\). This operator selects the set of propositions that constitutes a question and returns a (maximal) sum of propositions (“the most informative relevant answer”), i.e. a proposition, thus satisfying the selection requirement of resolutive predicates. Besides, in (39), as said in introduction, the DP the answer can substitute for the Wh-clause.

Lahiri (2002:92-94) challenged his view for Hintikka-like existential questions\(^{17}\). Still, there remain circumstances where no quantification takes place and where Chierchia’s operator is still required, though not intended for this purpose. These are the cases where no QVE arises, i.e. when a nonveridical operator is present with Wh-questions (see the discussion about (42) and (43), where the type the answer is still substitutable for the Wh-clause). Once again, this is also the context where yes/no UEQs are licensed. In this case the operator \(\sigma\) would select the only proposition that is the correct answer, depending on the context. In (4)a, it would be a member of the set \{Peter came; Peter did not come\}. The factive/veridical requirement of the main verb is satisfied by the fact that \(\sigma\) selects just the correct answers.

This proposal is supported by the parallel with concealed questions such as I (don’t) know the time. Concealed questions are definite (Nathan 2005) or specific indefinite DPs (Frana 2006, Faure 2010:chapter 6). Moreover concealed questions, though DPs, denote sets of propositions or propositions. Recall now that UEQ-clauses denote in fact true propositions. This true feature of the proposition denoted by the UEQ is in line with the definite or specific character of the concealed questions. That is why the assertive operator, sometimes identified with a D, as in Basque, and the definite/specific indefinite determiner in concealed questions could play the same role of type-shifter. Note that the extraction impossibility in English (see (19)) ensues as easily from the presence of a definite determiner as that of an NPI D.

Interestingly, Nathan 2005 defends the existence of two types of concealed questions, one of which is limited to resolutive predicates much as UEQs are. According to him, concealed question DPs undergo two type-shifting operations. The first one (44) turns the NP into a question, the second one (45) is operated by the determiner and turns the question into a proposition, much like Chiercha’s operator. Whether the operator is better described as \(\sigma\) or as \(\iota\), little matters here.

---

\(^{15}\) With the notable and often noted exception of speech act verbs like tell, cf. Égré 2008, Faure (2010:chapter 9) for tentative explanations.

\(^{16}\) Even in the case of UEQs, see footnote 5.

\(^{17}\) Like “Peter knows where to buy a book”, where the answer is felicitous even if it mentions only some (and not all) places where Peter can buy a book. Though it is still a topic of research, the Hintikka-like interpretation of questions could merely be a contextual restriction, the maximal sum being in this case reduced to a single element (on this point see Chierchia 1993:footnote 13).
(44) $\lambda P_{\langle s, t\rangle >.}\lambda p_{\langle s, t\rangle >.}[\exists x.e.p = \lambda w.x.P(w)(x)]$ 

$\langle s <e, t\rangle > \rightarrow \langle<s, t\rangle, t\rangle$

(45) $[[\text{the}]] = \lambda Q_{\langle s, t\rangle >.}\lambda p_{\langle s, t\rangle >.}[Q(p) = 1]$ 

(cf. $[[\text{the}]] = \lambda P_{\langle e, t\rangle >.}\lambda x.e.[P(x) = 1]$)

5.3. Summary

In this section I addressed the semantic side of the UEQ issue and tried to clarify the type-mismatch between proposition-selectors and questions. I argued against Adger and Quer’s IR approach to this issue and claimed that the problem is better solved by positing a type-shifting operation. This proposal is strengthened by its kinship with operations that are requested to explain $Wh$-UEQs in nonveridical contexts and concealed questions.

6. Concluding remarks

In this article, I studied the long-standing question of the distribution of embedded questions and the machinery needed to account for it, especially when there is both a semantic and a syntactic selection mismatch, i.e. in the case of Unselected Embedded Questions with resolutive (proposition-selecting) predicates. These clauses can all be substituted for a DP like the answer. They then denote a proposition, a denotation achieved through a type-shifting operation. The type-shifting operator takes in some languages the overt form of a D or an assertive C. It is covert in other languages, but detectable through extraction tests. At the interface with the selector, the UEQ is then a DP or an assertive CP and no syntactic mismatch arises.

The issue then boils down to the nonveridical character of the UEQ head. It is licensed by the distribution of yes/no UEQs which appear only in nonveridical contexts. As mentioned about examples (6)a/b, this does not seem to apply to $Wh$-UEQs. I have only sketched an answer to this problem. As seen in section 5.1, $Wh$-UEQs do not behave in the same way in veridical as in nonveridical contexts. In nonveridical contexts, $Wh$-UEQs pattern with yes/no UEQs in not allowing for quantification. This is not the case in veridical contexts. This point has to be taken into account and may indicate that we are not dealing with the same type of clauses as in veridical contexts, something that should be detectable through a more thorough examination of their syntactic properties.

References


Giannakidou A. (1998) Polarity sensitivity as (non) veridical dependency, John Benjamins,
Amsterdam-Philadelphie.


ASSOCIATION BETWEEN FOCUS PARTICLES AND INTERROGATIVE WH-PHRASES*

Haoze Li
The Chinese University of Hong Kong

1. Introduction

Aoun & Li (1993) discuss an interesting phenomenon—the association between focus particles and interrogative wh-phrases (F-WH association, henceforth) in Mandarin. Specifically, it refers to the phenomenon that in a wh-question, a focus particle must be associated with a wh-phrase, as in (1a-c). In these sentences, the wh-phrases occupy positions with which the focus particles are able to associate (details to be discussed in section 2) (in this paper, a focus particle and its focused associate are set in boldface and underlined, respectively).

(1) a. Libai zhi zai nali du shu?
   Libai only in where read book
   ‘What is the place x such that Libai reads books only in x?’

b. Libai zhiyou shenme cai cai gei nv’er zuo?
   Libai only what dish just for daughter cook
   ‘Which is the dish x such that Libai only cook x for his daughter?’

c. Shi shei zuotian qu-le Guangzhou?
   SHI who yesterday go-Asp Guangzhou
   ‘Who was the person x such that it is x who went to Guangzhou yesterday?’

Although F-WH association has been also noted in other studies (Shi 1994, Zhu 1997, Li 2011, Yang 2012), its importance to the theory of grammar has not been properly explored. This study takes up the theoretical consequences of F-WH association, showing that it not only challenges Beck’s (2006) General Minimality Effects (GME, henceforth), but also poses problems for three major approaches to wh-in-situ questions, namely the LF movement approach (Huang 1982), the unselective binding approach (Tsai 1994, 1999) and the choice function approach (Reinhart 1997, 2006). This study then proceeds to propose that wh-questions are best analysed along the lines of Hamblin’s (1973) classic compositional system, which makes it possible to reduce F-WH association to the phenomenon widely known as “Association with Focus” (AwF, henceforth) (Jackendoff 1972, Rooth 1985, 1992).

The paper is organized as follows. Section 2 describes the distributional and interpretive properties of F-WH association, highlighting the similarities between F-WH association and AwF. Sections 3 and 4 illustrate how F-WH association raises questions for Beck’s (2006) GME and previous approaches to wh-in-situ questions, respectively. I offer my proposal based on Hamblin (1973) in section 5 and conclude the paper in section 6.

* I would like to thank Candice Cheung, Lawrence Cheung, Chuansheng He, Jess Law, Thomas Lee, Jowang Lin, Zhuang Wu, Jiuhui Yang and Jing Yang for their valuable comments. I also thank the audience at GLOW in Asia IX, in particular, Jim Huang and Satoshi Tomioka, for their comments and insightful suggestions. I am responsible for all the remaining errors and shortcomings.
2. Similarities between F-WH Association and AwF

The current paper only includes the focus particles *shi*, *zhi* ‘only’ and *zhiyou* ‘only’, the focus marking function of which has been well studied. According to Aoun & Li (1993), Cheung (2008), Tsai (2004), Xu (2002, 2004), Zhang (1997, 2000) and Zhu (1997), these focus particles occur only in preverbal positions and must be associated with preverbal focused phrases they immediately c-command, as in (2)-(3).

\[(2)\]
\[
a. \text{Libai } zhi/\text{zhiyou } [\text{Xi’an}]_F \text{ qu nian } [VP \text{ qu-guo liang ci}].
\]
\[
\text{Libai } only \text{ only } \text{ Xi’an last year go-Asp two Cl}
\]
\[
\text{‘Libai visited only Xi’an twice last year.’}
\]
\[
b. *\text{Libai } zhi/\text{zhiyou } \text{Xi’an } [\text{qu nian}]_F [VP \text{ qu-guo liang ci}].
\]
\[
\text{Libai } only \text{ only } \text{ Xi’an last year go-Asp two Cl}
\]
\[
\text{‘Libai visited Xi’an only twice last year.’}
\]

\[(3)\]
\[
a. \text{SHI } [\text{Libai}]_F \text{ zuotian } [VP \text{ tan-le } \text{gangqin}].
\]
\[
\text{SHI Libai yesterday play-Asp piano}
\]
\[
\text{‘It is Libai who played the piano yesterday.’}
\]
\[
b. *\text{SHI } \text{Libai } [\text{zuotian}]_F [VP \text{ tan-le } \text{gangqin}].
\]
\[
\text{SHI Libai yesterday play-Asp piano}
\]
\[
\text{‘It is yesterday that Libai played the piano.’}
\]

In addition, a postverbal object can be associated with a preverbal focus particle and be marked as a contrastive focus, as in (4a-b).

\[(4)\]
\[
a. \text{Libai } zhi [VP \text{ song-le } [shu]_F \text{ gei Wangwei}].
\]
\[
\text{Libai only send-Asp book to Wangwei}
\]
\[
\text{‘Libai sent only books to Wangwei.’}
\]
\[
b. \text{Libai } \text{shi } [VP \text{ song-le } \text{shu } [\text{gei Wangwei}]_F ].
\]
\[
\text{SHI send-Asp book to Wangwei}
\]
\[
\text{‘It was Wangwei who Libai sent books to.’}
\]

If a focused constituent is outside the c-command domain of a focus particle, no AwF can take place, as shown in (5a-b).

\[(5)\]
\[
a. *[\text{Libai}]_F \text{ zhiyou/zhichi } \text{zai wangshang cai tan gangqin}.
\]
\[
\text{Only Libai played the piano at evening.’}
\]
\[
b. *[\text{Zhe ben shu}]_F, \text{shishi } \text{Libai mai-le}.
\]
\[
\text{‘It is this book that Libai bought.’}
\]

For F-WH association, *wh*-associates share the identical distributional properties of non-*wh* associates: (a) a focus particle must immediately c-command a preverbal *wh*-phrase, as in (6)-(7); (b) association can take place between a focus particle and a post-verbal *wh*-associate, as in (8a-b); (c) association cannot take place when a focus particle does not c-command a *wh*-phrase, as in (7a-b).
(6) a. Libai *zhi*/*zhiyou* na_ _ge difang qu nian [VP qu-guo liang ci]?
Libai only only which CI place last year go-Asp two CI
‘Which place is the x such that Libai visited only x in the last year?’

b. Libai *zhi*/*zhiyou* Xi’an _shenme shihou_ [VP qu-guo liang ci]?
Libai only only Xi’an what time go-Asp two CI
‘What is the time x such that Libai visited Xi’an twice only in x?’

(7) a. _Shi_ shei zuotian [VP tan-le gangqin]?
SHI who yesterday play-Asp piano
‘Who is the person x such that it is x who played the piano yesterday?’
b. *Shi_ Libai _shenme shihou_ [VP tan-le gangqin]?
SHI Libai what time play-Asp piano
‘What is the time x such that it is in x that Libai played the piano?’

(8) a. Libai _zhi_ [VP song-le _shenme_ gei Wangwei]?
Libai only send-Asp what to Wangwei
‘What is the x such that Libai only sent x to Wangwei?’
b. Libai _shi_ [VP song-le _shu_ gei _shei_]?
Libai SHI send-Asp book to who
‘What is the person x such that Libai sent books to x?’

(9) a. *Shei_ zhiyou/zhi_ zai wangshang cai tan gangqin?
who only only at evening just play piano
‘Who is the person x such that only x plays the piano in evening?’
b. *Shenme shu, Libai _shi_ mai-le?
what book Libai SHI buy-Asp
‘What is the book x such that it is x that Libai bought?’

According to Xu (2002), Tsai (2004) and Li (2011), focused phrases associated with _shi_, _zhi_ and _zhiyou_ are treated as contrastive focus, expressing exhaustivity, i.e. it specifies an exhaustive set of discourse entities for which a given proposition holds true and excludes other possibilities. Following È. Kiss (1998), they devise denial tests as diagnostics for exhaustivity. The conversations in (10) illustrate a denial test.

(10) Context: Libai visited Xi’an and Beijing.

Speaker A: Libai qu-guo [Xi’an],
Libai go-Asp Xi’an
‘Libai visited Xi’an.’

Speaker B: (#Bu dui,) ta hai qu-guo Beijing,
no right he also go-Asp Beijing
‘(#No,) he also visited Beijing.’

Speaker A’: Libai _zhi_ qu-guo [Xi’an],
Libai only go-Asp Xi’an
‘Libai only visited Xi’an.’

Speaker B’: Bu dui, ta hai qu-guo Beijing,
no right he also go-Asp Beijing
‘No, he also visited Beijing.’

In the context, Libai visited both Xi’an and Beijing. The focused phrases in Speaker A’s
utterance is not associated with a focus particle, hence does not express exhaustivity. In other words, this utterance does not indicate that Libai only visited Xi’an, excluding Beijing. This utterance cannot be judged as false in the given context. Hence, the denial in Speaker B’s utterance is inappropriate. In contrast, the focused phrase in Speaker A’s utterance is associated with the focus particle zhi, hence expresses exhaustivity. It means that the set of cities Libai visited only includes Xi’an and excludes Beijing. The utterance is false in the context. Therefore, the denial in Speaker B’s utterance is appropriate.

If association can take place between focus particles and wh-phrases, we should expect to attest exhaustivity in the wh-phrases associated with focus particles. Generally, if a wh-phrase bears exhaustivity, it requires its answer to share the same property (see Horvath 1986). The conversations in (11) illustrate how to diagnose exhaustivity in wh-phrases using the denial test.

(11) Context: Libai wants to drink wine and beer.

Speaker A: Libai xiang he shenme? Libai want drink what
‘What does Libai want to drink?’

Speaker B: Hongjiu. ‘Wine.’

Speaker C: (#Bu dui,) ta hai yao he pijiu. no right he also want drink beer
‘(#No,) he also wants to drink beer.’

Speaker A’: Libai shi xiang he shenme? Libai SHI want drink what
‘What is the thing x such that it is x that Libai want to drink?’

Speaker B’: Hongjiu. ‘Wine.’

Speaker C’: Bu dui, ta hai yao he pijiu. no right he also want drink beer
‘No, he also wants to drink beer.’

In the first conversation, the wh-question posed by Speaker A does not require a semantically exhaustive answer. Speaker B’s answer says that Libai wants to drink wine, and indeed he does. Since Speaker B’s answer is only incomplete, but not false, Speaker C cannot deny Speaker B’s answer. In contrast, in the second conversation, Speaker C’ can deny Speaker B’s answer. In Speaker A’s question, the wh-phrase associated with the focus particle shi semantically requires an answer with exhaustivity. Consequently, Speaker B’s answer describes a situation in which it is wine that Libai wants to drink, and he wants to drink nothing else. Obviously, it is false according to the context. Therefore, Speaker C’s response is appropriate in this situation.

In addition, we can use the partial truth diagnostics to test semantic exhaustivity in wh-questions with or without F-WH association. Consider the conversations in (12).

(12) Context: Libai wants to drink wine and beer.

Speaker A: Libai yao he shenme? Libai want drink what
‘What does Libai want to drink?’

Speaker B: Hongjiu. ‘Wine.’

Speaker C: Shi de, buguo ta ye yao he pijiu. yes DE but he also want drink beer
‘Yes, but he also wants to drink beer.’
Speaker A': Libai zhi yao he shenme? Speaker B': Hongjiu.
Libai only want drink what wine
‘What is the thing x such that it is x that Libai want to drink?’ ‘Wine.’
Speaker C': #Shi de, buguo ta ye yao he piju.
yes DE but he also want drink beer
‘Yes, but he also wants to drink beer.’

For example, in the first conversation in (12), Speaker C can credit Speaker B by affirming that his answer is partially correct. This is because although speaker B’s answer is incomplete, it is still true in this context. In contrast, in the second conversation, what Speaker A’ produces is a wh-question with F-WH association. It requires a semantically exhaustive answer. By answering “wine”, Speaker B’ really means that Libai only wants to drink wine and nothing else. Clearly, his answer is not only incomplete but also false. So, Speaker C’ cannot affirm that Speaker B’’s answer is partially true.

In sum, I have demonstrated that F-WH association shares distributional and interpretive properties with AwF. I take these similarities as evidence that F-WH association is a subtype of AwF. The possibility of F-WH association should not be surprising, given the vast amount of literature advocating that wh-phrases are focus (Horvath 1986, Rochemont 1986, Shen 1990, Zubizarreta 1998, Liu & Xu 2005). If a wh-phrase must function as the focus of a wh-question, it is expected that it can be associated with a focus particle. In section 6.2, I show that the compositional system of AwF is also applicable to F-WH association.

3. GME

Based on intervention effects in wh-questions, Beck (2006) proposes GME, the statement and the configuration of which are given in (13a) and (13b), respectively.

(13)

a. The evaluation of alternatives introduced by an XP cannot skip an intervening focus interpretation operator, ~ Op;
b. *[Op₁ …[ ~Op [ …XP₁ …]]]

Beck adopts the modified alternative semantics (Krater 1991, Wold 1996) to interpret AwF. In addition, she argues that a wh-phrase has a focus semantic value, denoting a set of alternatives, but its ordinary semantic value is undefined. In her analysis, the question operator (Q-Op, henceforth) in the CP layer is a kind of focus sensitive operator. It can evaluate a wh-phrase and define the ordinary semantic value of a wh-question.

Let me take one of Beck’s examples and concretely show how GME works. The example (14a) comes from Korean. In this sentence, the focused subject associated with the focus particle – man can introduce the ~Op into the LF representation, as in (14b).

(14)

a. *[Minsu-man]₇ nuku-lūl po-ass-ni?
Minsu-only who-Acc see-Past-Q
‘Who did only Minsu see?’ (Beck 2006: 1)
b. *[CP Q-Op [i₁ ~Op [ i₂ [Minsu-man]₇ nuku-lūl po-ass-ni]]]

Crucially, since wh-phrase has undefined ordinary semantic value, the ordinary semantic value of IP2 is also undefined. According to Rooth (1992), the ~Op must make use of both the ordinary semantic value and the focus semantic value. Due to the undefined ordinary semantic value of IP2, the composition between the ~Op and IP2 fails. As a result, IP1 is undefined. Furthermore, since the ~Op intervenes between the Q-Op and the wh-phrase,
GME prevents the Q-Op from evaluating the \( wh \)-phrase, leaving the ordinary semantic value of whole sentence undefined. According to Beck, the undefined ordinary semantic value leads to the ungrammaticality of the LF representation.

Beck’s original work has been well noted in the literature and has influenced some recent studies, such as Kim (2006), Beck & Kim (2006), Cable (2010) and Erlewine (to appear). Nevertheless, like many other general principles, GME runs into problem when a wider array of crosslinguistic data is taken into consideration. Specifically, I will show that in Mandarin a \( wh \)-question is perfect even if there is an intervening focus particle. Consider (15a) with the LF representation (15b).

(15) a. Libai zhi [VP hui du shenme shu]?
   Libai only will read what book
   ‘What is the book x such that Libai will read only x?’
   b. [CP Q-Op [IP …zhi [~Op [VP … wh-phrase]]]]

Obviously, (15b) is similar to the typical GME configuration in (13b), but it represents a fully grammatical sentence.

Moreover, Beck’s analysis leads to the prediction that a focus particle cannot be associated with a \( wh \)-phrase. Following Rooth (1985, 1992), I propose that the semantic interpretation of the focus particle zhi ‘only’ in (15a) is (16).

(16) \( \lambda x \forall p [p \in C \& \text{true}(p) \rightarrow p = \text{VP}'(x)] \)

If the ordinary semantic value of the \( wh \)-phrase were undefined, as claimed by Beck (2006), the ordinary semantic value of the VP would be undefined. As a result, the expression after ‘\( \rightarrow \)’ would also be undefined, i.e. without a truth value. Accordingly, (16) would be ill-formed. This means that F-WH association should be totally disallowed in natural languages. However, the fact that F-WH association is allowed in Mandarin, as shown in section 2, challenges this conclusion.

4. Challenges for Previous Approaches to \( Wh \)-questions

We have seen that F-WH association poses problems for Beck’s GME. In this section, I further show that F-WH association challenges three approaches to \( wh \)-questions in Mandarin, namely the LF movement approach, the unselective binding approach, and the choice function approach.

4.1 LF Movement

The classic analysis of LF movement originates from Chomsky (1977) and is developed by Huang (1982), who argues that an in-situ \( wh \)-phrase is a quantifier and undergoes covert movement to Spec-CP. According to this view, the LF representation of (17a) is (17b).

(17) a. Libai kanjian-le na ge ren?
   Libai see-Asp which Cl person
   ‘Which person did Libai see?’
   b. [CP [na ge ren]_i [IP Libai kanjian-le t_i]] (LF representation)

Although the LF movement approach has been influential, it faces some nontrivial problems, among which is F-WH association. As pointed out by Aoun & Li (1993), in (18a)
the LF movement of the in-situ *wh*-phrase results in the failure of F-WH association, as visualized in (18b).

(18) a. Libai  
  
  b. [CP  

In order for AwF to take place, the focus particle must c-command its associate (see section 2). Aoun & Li (1993), based on quantifier scope interactions and Antecedent Contained Deletion, argue that the c-command requirement must apply to the LF representation as well. In (18b), since *zhi* does not c-command the *wh*-phrase, F-WH association should not be possible, contrary to fact.

There is a possible solution to this problem. Since the focus particle is associated with the *wh*-phrase in (18a), the latter may be assumed to undergo covert movement to a position adjacent to the former, as in (19a). They may be combined together through reanalysis, and move further to Spec-CP, as in (19b). In this way, the *wh*-phrase moves to Spec-CP and the focus particle can still be associated with it.

(19) a. [CP  

However, this solution leads to new problems. If there were indeed covert movement of the *wh*-phrase, it should be constrained by island conditions. However, the sentences in (20) - (21) show that focus particles can be associated with *wh*-phrases inside islands.

(20) a. Libai  

b. [CP  

(21) a. Libai  

b. [CP  

Even though the island insensitivity can be handled by Fiengo et al. (1988), their solution leads to another problem on scope interactions. Consider the sentences in (22a-b).

(22) a. Libai  

b. Libai  

In (22a), the modal verb *keyi* ‘can’ scopes over the focus particle *zhi*, and the answer to the *wh*-phrase is a person such that Libai is allowed to only dance with this person. Under this interpretation, Libai is allowed to dance with others. In (22b), *zhi* scopes over *keyi*, and the answer to the *wh*-phrase is a person who is the only one partner Libai can dance with. Under this interpretation, Libai is not allowed to dance with anyone else.
According to the derivations in (19a-b), (22a) should have an LF representation in which the focus particle *zhi ‘only’ should have a wider scope than the modal verb *keyi ‘can’, as in (23), since the *wh-phrase could covertly move to the adjacent position to *zhi and the *zhi-*wh-phrase complex could undergo further LF movement to Spec-CP.

\[(23) [\text{CP} [\text{zhi} [\text{he shi}]]_2 [\text{IP} \text{Libai} \text{keyi} t_2 t_1 \text{tiaowu}]]\]

Obviously, *zhi should scope over the modal verb *keyi in (23). This means that (22a) should have the same interpretation as (22b), contrary to fact.

### 4.2 Unselective Binding

In order to resolve the problem of F-WH association, Aoun & Li (1993) proposes an in-situ binding approach, i.e., an in-situ *wh*-phrase does not undergo any movement and is bound by a Q-Op. This approach is developed by Tsai (1994, 1999), who makes use of Heim’s (1982) mechanism for interpreting indefinites. According to this approach, a Q-Op in Spec-CP can unselectively bind all the free variables in its scope. According to this approach, the LF representation for (24a) is roughly (24b).

\[(24) \begin{align*}
\text{a. Libai shenme difang qu-guo liangci?} \\
\text{Libai what place go-Asp twice} \\
\text{‘Which place did Libai visit twice?’}
\end{align*} \]

\[(24) \begin{align*}
\text{b. \text{[CP Q-Op}_1 \text{ [IP Libai [shenme difang]}_12 qu-guo liangci]}}
\end{align*} \]

Working under the unselective binding approach, Shi (1994) uses the operator binding mechanism to explain F-WH association: a *wh*-phrase is treated as an emphatic variable, which can be bound by emphatic operators. This is evidenced by the fact that the focus of a *wh*-question is the *wh*-phrase (see section 2). Both focus particles and the Q-Op are emphatic operators. In his account, one emphatic variable can be bound by two emphatic operators simultaneously. Thus, the LF representation of (25a) is (25b).

\[(25) \begin{align*}
\text{a. Libai shi shenme difang qu-guo liangci?} \\
\text{Libai SHI what place go-Asp twice} \\
\text{‘What is the place x such that it is x that Libai visited twice?’}
\end{align*} \]

\[(25) \begin{align*}
\text{b. \text{[CP Q-Op}_1 \text{ [IP Libai shi}_2 [shenme difang]}_121 qu-guo liangci]}
\end{align*} \]

It is clear that F-WH association is reduced to operator-variable binding. However, (25b) violates the locality condition of binding. In the logical representation, the focus particle, as an emphatic operator, is closer to the *wh*-phrase than the Q-Op. Hence, when the focus particle binds the *wh*-phrase, it is impossible for the Q-Op to bind the same *wh*-phrase.

Exploring the literature, the locality condition on operator binding might be made amenable to Lee’s (1986) modified Bijection Principle (I), which allows a variable to be bound by two quantifiers of the same type. Therefore, the *wh*-phrase in (25) can be bound by both the Q-Op and the focus particle, since the two binders are both emphatic operators. Nevertheless, this solution faces empirical problems. In Mandarin, one focused phrase cannot be associated with two focus particles simultaneously, as shown in (26a-b).

\[(26) \begin{align*}
\text{a. \text{**Zhiyou}_1 shi}_2 \text{ Libai}_12 qu-guo Xi’an.} \\
\text{only SHI Libai go-Asp Xi’an}
\end{align*} \]

\[(26) \begin{align*}
\text{b. \text{**Libai zhiyou}_1 zhi}_2 \text{ zhe ben shu}_12 \text{ du-guo.} \\
\text{Libai only only this Cl book read-Asp}
\end{align*} \]
According to Shi (1994), the focus particles are emphatic operators. If the modified Bijection Principle (I) could adequately account for F-WH association, it should license (26a-b), in which two emphatic operators bind one emphatic variable. Unfortunately (26a-b) are outright unacceptable.

Besides violating the locality condition on operator binding, analysing F-WH association under the unselective binding approach also raises a problem on semantic implication. As pointed out by Reinhart (2006), an unselectively bound in-situ wh-phrase is translated into a free variable with its N-restriction. For example, shei ‘who’ is translated into “person(x)”. Note that the N-restriction stays in situ, rather than occurring as a restriction on the operator. As a result, anything that does not belong to the set denoted by the N-restriction can be regarded as a felicitous answer to the wh-question. Sentence (27a) illustrates this point.

\[(27)\]
\[a.\text{ Ruguo ta } zhi \text{ qing na ge gangqingjia, Libai jiu hui hen shengqi?} \]
\[\text{if he only invite which Cl pianist Libai just will very angry}\]
\[‘What is the pianist } x \text{ such that if he only invites } x, \text{ Libai will be very angry?’}\]
\[b. \{p: \exists x [p = (he only reads x & pianist(x)) } \rightarrow \text{ (Libai will be very angry)]}\]

The logical representation of (27a) is (27b). It turns out that the value for } x \text{ can be anything in the world, since its N-restriction occurs in the antecedent clause of an implication. Suppose } Russell, \text{ a non-pianist, is chosen, the antecedent clause is false, and the implication must be true. So, an appropriate answer to this question would be ‘if he only invites } Russell, \text{ Libai will be very angry’. However, (27a) cannot be answered in this way.}

The representation yielding the set of felicitous answers should be one where the N-restriction is pulled out of the implication, as in (28). This representation correctly allows the values for } x \text{ to be all and only pianists.

\[(28)\]
\[\{p: \exists x [pianist(x) & p = (he only invites x} \rightarrow \text{ Libai will be angry]}\}

However, this goes against the spirit of the unselective binding approach. Since the wh-phrase na ge gangqingjia ‘which pianist’, as shown in (28), must be extracted out of the implication clause, the focus particle can no longer associate with it.

### 4.3 Choice Function

To resolve the implication problem mentioned in section 4.2, Reinhart (1997, 2006) proposes the choice function approach, which has been developed by Winter (1997) and Kratzer (1998). In this approach, the main idea of unselective binding is maintained, but the variable is of a different sort. In wh-in-situ questions, wh-phrases introduce variables over choice functions, instead of variables over individuals, and the Q-Op in Spec-CP, as an existential quantifier, binds the choice function variable.

According to Reinhart, a function is a choice function if it applies to any nonempty set and yields a member of that set.\(^1\) As a result, the choice function can guarantee that only members of a specified set are considered. According to this approach, the logical representation of (27a) is (29).

\[(29)\]
\[\{p: \exists / [CH(/) & p = (he only invites /pianist) } \rightarrow \text{ Libai will be angry]}\]

\(^1\) Although Winter (1997) and Kratzer (1998) revise Reinhart’s original choice function approach, they share the definition of the choice function with Reinhart.
(29) says that a function exists, such that if he invites the pianist selected by this function, Libai will be very angry. Here, the function cannot pick a member from sets other than the set of pianists.

However, the choice function approach cannot solve Beck’s (2006) problem. In (27a), the focus particle is associated with the *wh*-phrase. Under the choice function approach, the *wh*-phrase is translated as “f(pianist)” in the logical representation (29). Since f is a variable, f(pianist) also has an undefined ordinary semantic value. Consequently, the semantic composition does not allow association between the *wh*-phrase and the focus particle, wrongly ruling out F-WH association.

Furthermore, the choice function approach also raises a problem with the belief context. Consider the example in (30a-b).

(30)a. Libai xiangxin Wangwei *zhi* kanjian-le shei?
Libai believe Wangwei only see-Asp who
‘Who is the person x such that Libai believes that Wangwei only saw x?’

b. \{p: \exists f [CH(f) & p = Libai believe_w2 Wangwei only saw f(person)_w2 ]\}

In (30a), the focus particle is associated with the *wh*-phrase, and the association prevents the *wh*-phrase from being extracted out of the embedded clause. According to the logical representation (30b), since the *wh*-phrase must be scoped over by the verb “believe”, it should yield a person in Libai’s belief world rather than in the actual world. Consequently, this logical representation wrongly predicts that *wh*-phrases cannot have the de re reading.

5 Proposal

So far, we have seen that F-WH association causes problems for Beck’s (2006) GME as well as three well-known approaches to *wh*-questions. Drawing insights from Hamblin’s (1973) classic semantics of *wh*-questions, I propose that F-WH association can be accounted for by the same compositional system handling AwF.

5.1 Hamblin’s Semantics for *Wh*-questions

Since Hamblin (1973), the semantics of *wh*-questions has been standardly assumed to denote sets of propositions. Although this has been implemented in various ways by its successors (see Karttunen 1977, Hagstrom 1998, Reinhart 1997, 2006), I adopt Hamblin’s (1973) original work. Under his framework, a question simply characterizes a set of propositions, which serve as possible answers to the question. In other words, knowing the meaning of a question is equivalent to knowing what can be counted as answers. Given this, the answer set for (31a) would be the set in (31b).

(31)a. Libai piping shei?
Libai criticize who
‘Who does Libai criticize?’

b. \{Libai criticizes Bach, Libai criticizes Mozart, \…\}

According to Hamblin (1973), a *wh*-phrase like shei ‘who’ in (31a) denotes a set of alternative individuals, made up of the possible candidates for substituting the *wh*-phrase in the answer, as in (32a). In addition, a verb like piping ‘criticize’ in (31a) denotes a singleton set whose only member is its ordinary denotation, i.e., a function, as in (32b).
The composition of the verb and the wh-phrase is facilitated by the pointwise functional application rule, as stated in (33).

Pointwise functional application rule (Kratzer & Shimoyama’s (2002) version)
If \( \alpha \) is a branching node with daughters \( \beta \) and \( \gamma \), and \( [[\beta]]^{w,g} \subseteq D_{<s,c>} \) and \( [[\gamma]]^{w,g} \subseteq D_{<s,c>}, \) then \( \langle \alpha \rangle^{w,g} = \{ a \in D_s : \exists b \exists c [ b \in [[\beta]]^{w,g} \land c \in [[\gamma]]^{w,g} \land a = c(b) \} \}

Simply speaking, applying the function in (32b) to each member in (32a) results in a set of properties in (34).

Now we compose the set denoted by the VP piping shei with the proper name Libai, again by the pointwise functional application rule, and thus form a set of propositions in (35).

Finally, according to Kratzer & Shimoyama (2002), a Q-Op is applied to the set of propositions in (35). The Q-Op leaves the set intact, producing a question meaning, as in (36).

The above illustration is a sketchy description of how the interpretation of a wh-question is derived. In the process of the derivation, the pointwise functional application rule plays a critical role for obtaining the final representation. Note that the pointwise functional application rule is able to evaluate a wh-phrase in a displaced position though the wh-phrase is interpreted in situ (see Eckardt 2007, He 2011). This is achieved by the process of expansion inherent to the pointwise functional application rule, i.e., expanding sets of individuals to sets of properties, and from sets of properties to sets of propositions, as visualized in (37).
Section 5.3 will show that the expansion mechanism can help us resolve the island problems faced by the LF movement approach and the interpretative problems faced by the unselective binding approach.

5.2 Semantic Composition for F-WH Association

Based on Hamblin’s semantics of wh-questions, I illustrate how to interpret F-WH association. Consider the wh-question in (38). As presented in section 5.2, the VP piping shei ‘criticize who’ denotes a set of properties in (39).

(38) Libai zhi piping shei?
    Libai only criticize who
    ‘Who is the person x such that Libai only criticizes x?’

(39) \[[piping shei]\] = \{a: x [person(x)(w) & a = \lambda y\lambda w’[criticize(x)(y)(w’)]]\}
    = \{\lambda y.l.w’[criticize(Bach)(y)(w’)], \lambda y.l.w’[criticize(Mozart)(y)(w’)], \ldots\}

Now, we apply the focus particle zhi to the set of properties through the pointwise functional application rule, i.e., zhi is applied to each member of the set. As a result, we get a set in (40).

(40) \[[zhi]\]^{w,g}([[piping shei])]^{w,g}
    = \{[[zhi]]^{w,g}(\lambda y.l.w'[criticize(Bach)(y)(w')]),
        [[zhi]]^{w,g}(\lambda y.l.w'[criticize(Mozart)(y)(w')]), \ldots\}

Each member of the set in (40) provides an ordinary semantic value for zhi; and other members, as alternatives, provide a focus semantic value for zhi. This effectively solves the problem of undefined ordinary semantic value. Based on the denotation of zhi (see also Rooth 1985, 1992), we can rewrite (40) as (41).

(41) \{[[zhi]]^{w,g}(\lambda y.l.w'[criticize(Bach)(y)(w')])
    = \lambda y.l.w\forall p [p \in \{\lambda w'[criticize(Bach)(y)(w')], \lambda w'[criticize(Mozart)(y)(w')], \ldots\} &
        p(w) = 1 \rightarrow p = \lambda w''[criticize(Bach)(y)(w'')]
    \}

\{[[zhi]]^{w,g}(\lambda y.l.w'[criticize(Mozart)(y)(w')])
    = \lambda y.l.w\forall p [p \in \{\lambda w'[criticize(Bach)(y)(w')], \lambda w'[criticize(Mozart)(y)(w')], \ldots\} &
        p(w) = 1 \rightarrow p = \lambda w''[criticize(Bach)(y)(w'')]
    \}

\ldots

Apparently, the ordinary semantic value of the expression after “\rightarrow” in each member of the set is defined. Therefore, against the expectation of Beck’s (2006) analysis, F-WH association is possible.

Then, the subject Libai is composed with the set in (41), resulting in a set of propositions, as shown in (42).

(42) \[[Libai zhi piping shei]\]^{w,g} =
    \{\lambda w\forall p [p \in \{\lambda w'[criticize(Bach)(Libai)(w’)], \lambda w'[criticize(Mozart)(Libai)(w’)], \ldots\} &
        p(w) = 1 \rightarrow p = \lambda w''[criticize(Bach)(Libai)(w’’)]
    \}

\{\lambda w\forall p [p \in \{\lambda w'[criticize(Bach)(Libai)(w’)], \lambda w'[criticize(Mozart)(Libai)(w’)], \ldots\} &
        p(w) = 1 \rightarrow p = \lambda w''[criticize(Bach)(Libai)(w’’)]
    \}

\ldots
The Q-Op is applied to the set and the set is the meaning of the *wh*-question in (38). According to Hamblin’s system, each of propositions contained in the set is a possible answer to the *wh*-question. Suppose that the answer is “Libai only criticize Bach”. Since the focus particle is associated with *Bach*, the answer claims that the person Libai criticizes is Bach and only Bach, not Mozart and others. The exhaustivity (see section 2) are correctly encoded.

5.3 Advantages

In this section, I examine whether our proposal can account for F-WH association adequately, i.e., accounting for F-WH association without getting trapped in the problems reviewed in section 4.

Crucially, in our proposal, F-WH association is reduced to association between focus particles and members of a set expanded from a *wh*-phrase, as shown in (40)-(41). As a result, F-WH association can be treated as a set of AwFs. Since AwF is allowed in natural languages, F-WH association is certainly allowed in languages.

Moreover, as shown in section 2, F-WH association has similar distributive and interpretative properties as AwF. According to our proposal, F-WH association can be reduced to AwF. Therefore, the similarity is expected.

Furthermore, since F-WH association is reduced to AwF, it is reasonable that F-WH association is interpreted by the same mechanism that interprets AwF. Hence, it is unnecessary to assume additional rules or principles but only make use of the basic concepts and mechanisms defined in Hamblin’s semantics and Rooth’s alternative semantics to interpret F-WH association.

In addition, I do not argue that a *wh*-phrase is inherently quantificational, so LF movement of *wh*-phrases is not needed. As a result, the c-command requirement of F-WH association is trivially satisfied in overt syntax and LF. Moreover, since there is no movement, the island insensitivity problem faced by the LF movement approach is resolved directly. The scope problem is also resolved, given that there is neither movement of a focus particle nor a *wh*-phrase out of the scope bearing modal verb, as shown in (22)-(23).

The current proposal is also free from the binding problems faced by the unselective binding approach. Unlike Shi’s (1994) analysis, F-WH association in our compositional system does not block the function of the Q-Op. The Q-Op takes as its argument the set of propositions expanded from a *wh*-phrase, whereas a focus particle is associated with each member of a set expanded from a *wh*-phrase, leaving the expansion intact, as shown in section 5.2. Consequently, F-WH association does not close the expansion so the Q-Op can still be applied to a set of propositions.

Although a *wh*-phrase stays in situ at LF, it must be interpreted in a different position at the interpretive level. As shown in (37), the *wh*-phrase, which denotes a set, undergoes proliferation and takes a wide scope. In other words, the *wh*-phrase must be semantically extracted from the formula. Note that the extraction does not result from syntactic movement but rather from the pointwise functional application rule. In this respect, the extraction is purely semantic, instead of syntactic. For F-WH association, according to our proposal, a focus particle does not close the expansion of a *wh*-phrase, i.e., the set denoted by the *wh*-phrase also undergoes proliferation. Therefore, the bottom-up composition of (38) looks as follows:

= {‘Libai only criticizes Bach, (not Mozart, …)’,
‘Libai only criticizes Mozart, (not Bach, …)’, …}
(43) \[[IP]\]^{w,g} = \{p: \exists x[person(x)(w) \& p = \lambda w'\forall p[p\in\{\exists z\in Alt(x) \& \{Libai criticize z\}\}] \rightarrow p = \lambda w'[\text{criticize}(x)(Libai)(w'')]\}

\[[Libai]\]^{w,g} = \{f: \exists x[person(x)(w) \& f = \lambda y\lambda w'\forall p[p\in\{\exists z\in Alt(x) \& \{y criticize z\}\}] \rightarrow p = \lambda w'[\text{criticize}(y)(z)(w'')]\}

\[[zhi]\]^{w,g} = \{f: \exists x[person(x)(w) \& f = \lambda Q\lambda y\lambda w'\forall p[p\in\{\exists z\in C \& p(w') = 1\} \rightarrow p = Q(y)]\}

\[[VP]\]^{w,g} = \{f: \exists x[person(x)(w) \& f = \lambda y\lambda w'[\text{criticize}(x)(y)(w''')]\}

\[[vpiping]\]^{w,g} = \{x: person(x)(w)\}

\[[shei]\]^{w,g} = \{x: person(x)(w)\}

Based on this, our proposal can avoid the implication problem discussed in section 4.2 and the problem with the belief context discussed in section 4.3. Let’s consider the sentence (27), repeated as (44a), in which F-WH association is embedded in the antecedent clause of the implication. By the pointwise functional application rule, the set denoted by the wh-phrase undergoes proliferation and is extracted out of the implication, as in (44b) (see also He 2011). Apparently, this logical representation does not trigger the implication problem.

(44) a. Ruguo ta zhi qing na ge gansongjia, Libai jiu hui hen shengqi?
   ‘What is the pianist x such that if he only invites x, Libai will be very angry?’
   b. \{p: \exists x[\text{pianist}(x) \& p = (he only invites x) \rightarrow (Libai will be very angry)]\}

Then, consider the sentence (30), repeated as (45a). In this sentence, F-WH association is contained in the clausal argument of the intensional verb xiangxin ‘believe’. According to the pointwise functional application rule, the set denoted by the wh-phrase is extracted and takes a scope wider than the intensional verb xiangxin ‘believe’, as shown in (45b) (see also He 2011). The logical representation correctly allows the de re reading of the wh-phrase.

(45) a. Libai xiangxin Wangwei zhi xihuan shei?
   ‘What is the person x such that Libai believes that Wangwei only like x?’
   b. \{p: \exists x[person(x)(w) \& p = Libai believe w'\forall Wangwei only like x in w']\}

6 Conclusion

This paper studies F-WH association in Mandarin, which is pointed out by Aoun & Li (1993). Empirically, I find that it shares the distributional and interpretive properties of AwF in Mandarin. Theoretically, based on Hamblin’s semantics of wh-questions, I propose that F-WH association can be reduced to AwF. The current study may shed light on the interpretive mechanism of in-situ wh-questions. Our proposal accounts for the availability of F-WH association in natural languages, a phenomenon wrongly excluded by Beck’s (2006) GME, and avoids many interpretive problems faced by previous approaches to wh-questions. If this study is on the right track, then Hamblin’s semantics may be the best tool for interpreting in-situ wh-questions.

References

THE SYNTACTIC STRUCTURES OF THE CHINESE BEI PASSIVE AND THE ENGLISH BE-PASSIVE*

Na Liu
Tianjin Normal University

1. Introduction

There are two types of Chinese bei passives – the long passive and the short passive, as shown in (1-2) respectively, depending on whether the Agent phrase is present or not.

(1) The long passive: bei NP-VP
   Zhangsan bei Lisi da le.
   ‘Zhangsan was hit by Lisi.’

(2) The short passive: bei VP
   Zhangsan bei da le.
   ‘Zhangsan was hit.’

A significant word order difference between the English be-passive and the Chinese bei passive is that in English, the external argument occurs at the end of the sentence as in John was hit by Bill, while in Chinese, the external argument precedes the main verb as in (1). From the traditional view that took the agent phrase (by- or bei-phrase) to be an adjunct, the word order difference simply follows as a special case of the difference in the location of adjuncts between these languages. However, the recent research has shown that the external argument actually stays in its base position, for both English (Collins 2005, Gehrke and Grillo 2009) and Chinese (Huang 2011, Huang & Liu 2011). Given the basic SVO word order for both languages, the difference in surface word order between their passives needs an explanation.

Considering the word order of the English be passive, Collins (2005) offers a smuggling approach. Gehrke and Grillo (2009, G&G hereinafter) provide a similar VP movement approach with a different motivation for the movement. In this paper, I will argue that the Chinese bei passive does not have VP movement under either Collins’ or G&G’s account, and this explains the word-order difference. I also show how this difference arises the way it does between the two languages.

---

* Some ideas in this paper come from my Ph.D. dissertation written in Beijing Language & Culture University. This paper has benefited particularly from extensive discussions with C.-T. James Huang. I am also indebted to Chris Collins, Mamoru Saito, William Snyder, Uli Sauerland, Grant Goodall, Dominique Sportiche, Michael Barrie, Chunyan Ning, Gang Gu, Jilun Lu, Caimei Yang for their valuable suggestions and comments. I am solely responsible for all the remaining errors and inadequacies.

1 Abbreviations used in this paper are as follows: CL: classifier; Exp: experiential aspect marker; FP: focus phrase; PERF: perfective aspect marker.
2. Reanalysis of the Structure of Chinese Long and Short Passives

2.1. The Previous Analyses of the Chinese Long and Short Passives

According to Feng (1995), Ting (1995, 1996), Huang (1999) and Huang, Li & Li (2009, HLL hereinafter), the Chinese long *bei* passive can be analyzed as a structure of complex predication, as shown in (3).

(3) Zhangsan \(\overline{bei} [\overline{IP} \overline{NOP}_i [\overline{IP} Lisi da le t_i]].\)
\[\text{Zhangsan BEI Lisi hit PERF}\]
‘Zhangsan was hit by Lisi.’

In (3), the main verb *bei* selects an active IP as its complement, within which a null operator (NOP) moves from the object position of the base verb and adjoins to IP. The main verb *bei* and the complement IP form a complex predicate, which selects the matrix subject as its single argument. The moved NOP object is bound by the matrix subject under predication.

Huang (1999) and HLL (2009) propose a control analysis for the short *bei* passive, as shown in (4).

(4) Lisi \(\overline{bei} [\overline{VP} \overline{PRO}_i da-le t_i]\)
\[\text{Lisi BEI hit-PERF}\]
‘Lisi was hit’

In (4), *bei* is a light verb, selecting an Experiencer subject and a VP complement, in which the Theme object is an empty category PRO which moves into Spec, VP and is controlled by the matrix subject.

In summary, both long and short *bei* passives are covered by the control/predication analysis, which, according to Huang (1999) and HLL (2009), is motivated by the evidence that the subject of *bei* may take subject-oriented adverbs, such as *guyi* ‘intentionally’ in (5).

Hence, it is base-generated as the subject of *bei* and assigned an Experiencer theta-role by it.

(5) Zhangsan *guyi* \(\overline{bei} (Lisi) da le.\)
\[\text{Zhangsan intentionally BEI Lisi hit PERF}\]
‘Zhangsan intentionally got hit by Lisi.’

The difference between the long and short passives is that the former involves NOP movement which is an A’-movement. This is evidenced by the following facts:

Firstly, Chinese long passives exhibit long-distance dependency, as shown in (6).

(6) Zhangsan \(\overline{bei} \overline{pai} jingcha zhua-zou \overline{le}.\) (HLL 2009: 125)
\[\text{Zhangsan BEI police arrest-away PERF}\]
‘Zhangsan was “sent-police-to-arrest” by Lisi.’

In (6), Zhangsan was arrested by the police, but Lisi was the person who sent the police to arrest him. Hence, there are two embedded clauses. The *pai* ‘send’ clause embedded under *bei* with Lisi being its subject, and the *zhua* ‘arrest’ clause embedded under *pai* ‘send’ with the police being its subject. This structure of the long *bei* passive is akin to English tough sentences (Feng 1995, Ting 1995, 1996, Huang 1999, HLL 2009). Chomsky (1981) analyzes tough construction as involving NOP movement and predication, as shown below.
The Syntactic Structures of the Chinese *Bei* Passive and the English *Be*-passive (N. Liu)

(7) This problem is too easy [CP NOP₁ for me to ask the teacher to help me solve t₁].

Secondly, Chinese long-distance passives exhibit island effects and allow the occurrence of resumptive pronouns. These are the diagnoses for A’-movement.

(8) (HLL 2009: 125)
Zhangsan bei wo tongzhi Lisi ba zanmei *(ta) de shu dou mai-zou le.*
Zhangsan BEI me inform Lisi BA praise him DE book all buy-away PERF
‘Zhangsan had me inform Lisi to buy up all the books that praise [him].’

The ungrammaticality of (8) can be explained if the object of the verb *zanmei* ‘praise’ in the relative clause modifying ‘book’ is assumed to undergo A’-movement which is blocked due to the violation of the Complex NP constraint proposed by Ross (1967). The fact that the derivation can be saved with a resumptive pronoun in that position further supports the A’-movement hypothesis because such a pronoun always appears in an A’-bound position.

Thirdly, Huang (1999) cited the following example from Chiu (1995) to support the A’-movement hypothesis in the complement clause of long passives. Chiu strongly argues that the particle *suo* is only involved in long passives and relative clauses, and it is triggered by the existence of wh-movement in both cases.

(9) zhexie shiqing bu neng bei tamen suo liaojie. (HLL 2009: 126)
these thing not can BEI they SUO understand
‘These things cannot be understood by them.’

Despite the extensive evidence for a control/predication analysis, more recent research has shown that a raising analysis should also be made available for certain *bei* passives, according to which the surface subject of *bei* is not base-generated, but is derived by movement.

### 2.2. The Possibility of Raising Analysis for *Bei* Passives

Huang (2011) and Huang & Liu (2011) observe that both short and local long *bei* passives allow idiom-chunks to be fronted under passivization, as in (10). Such examples imply a raising analysis for the subject of *bei*:

(10) pianyi dou bei (ta) zhan-guang-le
advantage all BEI he take-away-PERF
‘All the advantage was taken by him.’

In the case of short and local long passives without any subject-oriented adverbs, logic does not preclude a raising analysis, either. Hence such sentences as (11) may have to allow either a raising or a control/predication analysis:

(11) shu bei (Lisi) si-po-le.
book BEI Lisi tear-break-PERF
‘The book got torn by Lisi.’

Based on these and other considerations, Huang (2011) and Liu (2011) concluded that:

A: Raising analysis is required for short and local long passives with idiom-chunk subjects;
B: Raising analysis is possible for short and local long passives with no subject-oriented adverbs;
C: The following types of passive continue to require a control or predication analysis:
   a) In the short and long passives with subject-oriented adverbs;
   b) In the long distance passives that exhibit A’-bar movement properties.

3. The Newly Proposed Structures of Short and Local Long Bei Passives

In this section, I will consider new structures for the short and local long passives, like (11), with both control and raising possibilities. They have similar underlying structures with the English be passive proposed by Collins’ (2005). In Section 3.1, I will introduce Collins’ smuggling approach to the English be passive. In Section 3.2 & 3.3, I will argue that the smuggling operation does not apply for Chinese passives and propose the structures for the Chinese short and local long bei passives.


According to Collins (2005), the problem of the traditional analysis of the be-passive under the Government and Binding (GB) theory of the framework of Principles and Parameters (P&P) is that the external argument is generated in Spec, TP in the active, but in a completely different position, i.e. complement of the preposition by, in the passive. This is a violation of UTAH (Uniformity of Theta-Assignment Hypothesis) (Baker 1988:46, 1997:74). In the spirit of UTAH, the external θ-role should be assigned in the passive in the same way as in the active. To solve this problem, Collins proposes a smuggling approach to the English be passive. The derivation of (12a) is illustrated in (12b).

(12) a. The book was written by John.

As shown in (12b), the external argument is merged in Spec, vP in the same way as in the active. By in the agentive by-phrase is the head of the VoiceP and takes a vP as its complement. If the internal argument the book moves to Spec, IP directly, a wrong word order will be produced, as in the ungrammatical sentence *The book was by John written. In addition, such a movement crossing the external argument violates Minimal Link Condition.
(Chomsky 1995) or Relativized Minimality (Rizzi 1990, 2000). Therefore, movement of the internal argument *the book* to Spec, IP position is carried out in two steps: first, the participle phrase (PartP) *written the book* is moved to Spec, Voice, and then *the book* is moved out of the PartP to Spec, IP. The two-step process effectively smuggles the internal argument to Spec, IP position without violating minimality conditions.

I assume that Chinese short and local long passives also involve a VoiceP which is headed by the passive marker *bei*, though the displacement of the internal argument to Spec, IP does not involve smuggling, thus keeping the VP in its base position following the external argument.

### 3.2. No Smuggling Occurs in the Chinese *Bei* Passive

The word order difference between English and Chinese passives seems to imply that smuggling does not occur in Chinese. I assume that the absence of smuggling in Chinese is related to another property of Chinese, i.e., Chinese allows object to be preposed to a post subject but preverbal focus position (Shyu 1995), as illustrated in (13). However, such movement is not available in English.

\[(13) \quad \begin{align*}
\text{(a)} & \text{ Lisi kan-guo le naben shu} \quad \text{(Shyu 1995:100)} \\
& \text{Lisi read-Exp PERF that-CL book} \\
& \text{‘Lisi has read that book.’} \\
\text{(b)} & \text{ Lisi naben shu kan-guo le ti} \\
& \text{Lisi that-CL book read-Exp PERF} \\
\end{align*} \]

I assume the derivations of (13a,b) are as in (14).

\[(14) \]

As for (13a), no object preposing occurs. V selects its complement DP *na-ben shu* ‘that book’. V moves to v and the external argument *Lisi* which is base-generated in Spec, vP moves to Spec, TP for reason of Case. To derive (13b), the object *na-ben shu* ‘that book’ is preposed to the Spec of a focus phrase (FP), which is above vP but below TP. If the object moves from the complement position of V to Spec, FP directly, crossing the Agent argument merged in
Spec, vP, Relativized Minimality (RM) will be violated. Following Ura (2000), I assume that v has some strong nominal feature (Case feature or scrambling feature (Grewendorf & Sabel 1999)) that requires the internal argument to move to inner Spec, vP. From there it moves to Spec, FP. Such movement does not fully cross the external argument. Therefore, no violation of RM occurs. In addition, based on Shyu’s (1995) argument that object preposing in (13b) is an A-movement, I deduce that the whole movement chain must be an A-chain, and every intermediate landing site including the inner Spec, vP is an A-position.

In summary, I propose that since object can be preposed to an A-position above vP and below TP in Chinese passives, there exists an extra Spec, vP which can serve as an intermediate landing site for the object to take on its way of moving to Spec, TP without violating minimality conditions. Hence, there is no need for the less economic smuggling operation to apply, which requires more things to move together with the object. On the other hand, English does not have such kind of object preposing, the extra Spec, vP position is not available in the passive structure. In order to avoid violation of minimality, smuggling has to apply, as a last resort.

3.3. The Derivations of Chinese Short and Local Long Passives

According to Alexiadou (2005), passive sentences have been classified into two kinds: canonical passives and non-canonical passives, typically manifested by the English be and get passives respectively. Huang (1999), HLL (2009) and Liu (2012) observe that Chinese bei passives are akin to English get passives but different from English be passives. Hence, Liu (2012) classifies Chinese bei passives as non-canonical passives. In order to explain why non-canonical passives can be analyzed as either control or raising structures while canonical passives cannot, Huang (2012) assumes that non-canonical passives are formed by superimposing on the core predicate a semi-lexical verb, such as get and bei, whose meaning may include one or more points in the causative-unaccusative spectrum as shown in (15).

(15) THE CAUSATIVE-UNACCUSATIVE CONTINUUM:
cause > let > witness > undergo > be affected by > become > exist > be

He further proposes that the semi-lexical verbs get and bei may occupy more than one point in the continuum in (15). When occupying the point ‘undergo’, they select an Experiencer subject and form control structures. When being located at the point of ‘become’, they do not introduce any thematic subject and form raising structures. Furthermore, Liu (2012) supposes that both get and bei can be decomposed into EXPERIENCE (Exp) and BECOME (Bec),

which are projected in syntax as the Experience phrase (ExpP) and the Become phrase (BecP). VoiceP is embedded under BecP which is embedded under ExpP. In the get passive, the VoiceP is headed by by in the agentive by phrase, but in the bei passive, I assume what heads the VoiceP is bei. Such an assumption can help settle a long-existing argument among

---

2 Huang (1999) implies that the Experiencer subject of the get-passive is introduced by the Become head (Bec). The causative-ergative alternation such as Mary got John blamed for the mistake vs. John got blamed for the mistake depends on whether there is a CauseP layer in the derivation. Richards (2001) and Harley (2002, 2004) assume that get in the get+DP structure can be decomposed into a light verb BECOME in its semantics. Orfitelli (2011) assumes that all get-constructions include the light verb vbecome. Along with them, we assume that get in the get-passive contains a Become (Bec) component. Different from them, we make a finer analysis in which get also involves an Experience (Exp) component, and the Experincer subject is introduced by Exp. Considering bei in the bei passive, since the bei passive has much resemblance to the get passive in meaning, and get and bei are both grammaticalized from a causative verb, denoting ‘to get’ or ‘to receive’ (Liu 2012), we assume that bei can also be decomposed into EXPERIENCE and BECOME.
Chinese linguists about whether *bei* is a verb or a preposition: *bei* is base-generated in Voice, functioning like the English *by* which checks the Accusative Case feature of the external argument base-generated in the Spec of its complement vP (Collins 2005). Then, *bei* moves from Voice to Bec and Exp which are verbal heads. Therefore, *bei* seems to have the properties of both a verb and a preposition.

Based on the above arguments, I propose that the raising and control structures of the local long *bei* passive (1), repeated as (16), are shown in (17) and (18) respectively.

(16) Zhangsan bei Lisi da le.
    ‘Zhangsan was hit by Lisi.’

(17) The tree (17) shows the derivation of the raising structure of (16). To begin the derivation, the internal argument *Zhangsan* is based-generated as the object of the base verb *da ‘hit’*. Due to the lack of Case, it has to move. It first moves to inner Spec, vP, an available intermediate landing site, given our derivation of (13b) in (14). The external argument is merged in outer Spec, vP. The Voice head *bei* is merged with vP and checks its accusative Case feature with the external argument *Lisi* which is in outer Spec, vP and closer to it, and then moves to Bec. The internal argument moves to Spec, VoiceP and then to Spec, BecP and Spec, TP to get Nominative Case. There is no violation of minimality and hence smuggling is not needed.
The tree (18) shows the control structure of (16), in which the passive verb *bei* is base-merged in Voice and moved to Bec, head of BecomeP, and Exp, head of Experience phrase. I assume that in the control structure, *bei* is decomposed into an Experience, a Become, and a Voice component. The derivation is carried out in the following way. First, the verb *da-hit* selects a null pronominal PRO as its internal argument to form VP. PRO must be vacated from the object position for reason that it cannot occur in a governed position. It moves to inner Spec, vP, Spec, VoiceP and Spec, BecP. The Exp head introduces the Experiencer argument Zhangsan in its specifier, which finally moves to Spec, TP to get nominative Case and controls PRO which is in Spec, BecP. For the case involving control, there is no movement of an overt object beyond *bei*. No minimality issues arise, and smuggling is not needed.

The derivations of short (agentless) *bei* passive under raising and control analyses are similar to that of the local long passives in (17-18), except that the external argument is null.

### 3.4. More Evidence for the Lack of Smuggling in Chinese *Bei* Passives

In this section, I will present two tests – the quantifier floating test and “by-DP” constituency test to show the absence of smuggling operations in Chinese *bei* passives.

#### 3.4.1. The Quantifier Floating Test

Quantifiers immediately precede the DPs they quantify and can be stranded after the DPs move. The distributions of the floated quantifiers in Chinese *bei* and English *be* passives as in (19-20) respectively can be explained if we assume that Chinese does not have smuggling
while English does.

(19) a. pingguo bei Lisi quanbu mai-zou-le
   apple BEI Lisi all buy-away-PERF
   ‘The apples were all bought by Lisi.’
   b. pingguo quanbu bei Lisi mai-zou-le.
      Apple all BEI Lisi buy-away-PERF
      ‘The apples were all bought by Lisi.’

(20) a. *They were arrested by the police all.
   b. They were all arrested by the police.

In Chinese bei passives, the floated quantifier quanbu ‘all’ appears below the external argument, as in (19a), or above ‘bei+external argument’, as in (19b). However, in English be passives, all can only float between auxiliary and the participle but not below the agentive by-phrase, as in (20a, b). I assume that the derivations of (19a, b) are illustrated in (21a, b) respectively, and that of (20b) is in (22). In (21a), the quantifier phrase quanbu pingguo ‘all apples’ is based-generated in the complement position of V and moves first to inner Spec, vP, where the quantifier quanbu ‘all’ is floated, with the DP pingguo ‘apples’ it quantifies moving on its own to Spec, VoiceP and then to Spec, TP. In (21b), quanbu ‘all’ is floated in the Spec, VoiceP position. The positions where the floated quantifier can appear show that on its way to move to Spec, TP, the object has moved through these positions and left traces.

(21) a. 

```
TP
   DP
      pingguo[T]
         ‘apple’ T
         VoiceP
            t_i
            Voice’
            vP
              Voice bei
              vP
                 Lisi
                      vP
                         [quanbu[t_j]]
                             ‘all’
                               v
                                  v’
                                     V’
                                        V
                                           V
                                              V
                                                 V
                                                    V
                                                      V
                                                        V
                                                          V
                                                            V
                                                              V
                                                                V
                                                                  V
                                                                    V
                                                                      V
                                                                        V
                                                                          V
                                                                            V
                                                                             V
                                                                               V
                                                                                  V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
                                                                                      V
```
The above tree (22) shows the derivation of the English passive (20b), in which the PartP smuggles the object to Spec, VoiceP, a position above the Agent by-phrase. That’s why all cannot float below the external argument. Again, the difference comes from the fact Chinese allows object preposing, while English does not.\(^3\)

\(^3\) It can be noticed that the quantifiers cannot be floated in the positions where the internal argument is base-generated in both English and Chinese, as in (i) and (ii).

(i) *They were arrested all by the police.

(ii) * pingguo bei Lisi mai-zou-le quanbu apple BEI Lisi buy-away-PERF all
3.4.2. The “By-Phrase” Constituency Test

According to Huang (1999) and HLL (2009), unlike the by-phrase in English passives, the string of bei-DP does not behave as a constituent, as shown in (23).

   Zhangsan yesterday BEI Lisi hit-PERF
   (cf. John was hit by Bill yesterday.)

   b. *Zhangsan bei Lisi zuotian da-le.
      Zhangsan BEI Lisi yesterday hit-PERF
      (cf. John was hit yesterday by Bill.)

   c. *bei Lisi Zhangsan zuotian da-le.
      BEI Lisi Zhangsan yesterday hit-PERF
      (cf. It was by Bill that John was hit yesterday.)

(23b) and (23c) show that unlike the by-phrase by Bill in the English translation, bei-DP cannot move as a constituent across a time phrase or prepose to a sentence initial position. These facts can be explained if we assume that the English passive has smuggling while the Chinese one does not. As shown in (22), after the movement of PartP, the rest of VoiceP which includes by, the Agent DP the police and the trace of PartP, is a constituent. The trace of PartP is silent. That’s why the by-DP behaves like a constituent. However, since there is no smuggling in the Chinese passive, as shown in (17), bei, the head of VoiceP cannot form a constituent with the Spec of vP.

There seems to be a counter-example to this explanation given by Shi & Hu (2005), quoted from Chen (2001), as shown in (24).

(24) yihuir, zhe meimiao de shengyin bei shu, bei cao, bei (Chen 2001)
a while this beautiful DE voice BEI trees BEI grass BEI
yi-ge guangmo de kongjian tuoshi-le
one-CL wild DE space swallow-PERF
‘Not for a while, this beautiful voice got swallowed by trees, grass and a wild space.’

In (24), bei and the Agent DP are coordinated with the element of the same type, which seems to show that bei-DP is a constituent. However, this example should not be treated as a coordination test for the constituent status of “bei+DP”. It is a phenomenon of right node raising (RNR), as argued by Huang (1999) and Xiong (2010). It is similar to (25) which is a typical case of RNR:

(25) [John loves and Mary hates] oysters.

According to Huang (1999:7), the function of RNR is to identify whether the raised rightmost part (oysters) is a constituent or not, but not that of the remnant (John loves or Mary hates).

In this section, the extra evidence from the quantifier floating test and “by-DP” constituency test supports the hypothesis that smuggling operations occur in English passives but not in Chinese ones.

‘The apples were all bought by Lisi.’
We treat this as a case of the generalized phenomenon mentioned by Bošković (2004:685) that “Quantifiers cannot be floated in θ-positions”. The quantifier all and quanbu are in the complement positions (i.e. θ-positions) of the verbs arrested in (i) and mai-zou-le ‘buy-away-PERF’ in (ii).
4. On the Structure of Chinese Passives from G&G’s Perspective

In the next section, I will show that under G&G’s approach, the VP movement (or smuggling) is not possible in Chinese passives either.

4.1. Gehrke and Grillo’s (G&G 2009) Analysis

G&G (2009) pointed out the problem of Collins’ approach, i.e. it is hard to tell the limits of the smuggling operation. According to them, the motivation of the movement of VP to Spec, VoiceP is not to smuggle the internal argument crossing the external one but to promote a result state subevent to meet the temporal requirement of event structure. The details of their assumption are shown in (26).

(26)         AspP
              |   
              | Asp-T
    Asp’     |   
    Asp     |   
    VoiceP  |   
              |   
              | EVT-T
    Voice’   |   
    Voice    |   
    VP₁      |   
              |   
    DPₑₓᵗ   |   
    V₁’      
    V₁       
              |   
    VP₂      |   
              |   
    DPᵢₙᵗ   |   
    V₂’      
    V₂       
              |   
              | (XP)

Based on the theory of decomposition of event types, they assume that a BECOME component is associated with the lower VP-shell, i.e. VP₂ in (26), which introduces the internal argument DPᵢₙᵗ. VP₁ is associated with the component CAUSE and introduces the external argument DPₑₓᵗ. Getting insights from Demirdache and Uribe-Etxebarria (2000), G&G propose that in passives, Voice has the responsibility for grounding the event time (EVT-T) which is anchored within the resultative state subevent. This feature of Voice triggers the movement of VP2 which is associated with BECOME component to Spec, VoiceP to anchor the event time so as to make VoiceP available to the next temporal domain of the clause, i.e. the Aspect domain, where the assertion time (ASST-T) is located.

One piece of evidence for the decomposition of the BECOME component in the lower VP predicate given by G&G is that only those transitive verbs associated with some kind of resultative meaning (or involving BECOME) can form passives. Those verbs associated with simple event structures cannot, as shown in (27).

(27) a. This laptop weighed two kilos. (G&G 2009:11)
    b. *Two kilos were weighed (by this laptop).

4.2. The Non-VP Movement Analysis of Chinese Passives under G&G’s Approach

Under G&G’s (2009) account, movement of the lower VP shell, i.e. VP₂ in (26), is to
convey the BECOME component associated with VP₂ to Spec, VoiceP where it can anchor the Event Time and make it available to operate with the Assertion Time in AspP. However, in Chinese, as assumed in Section 3.3, the passive verb bei can be decomposed into a BECOME component. I assume that this BECOME is projected into a Become phrase (BecP) between AspP and VoiceP as in (28). The event time is located in this Become phrase and is associated with the Assertion time in AspP. Hence, there is no need for the movement of the lower VP to Spec, VoiceP. However, in the be passive, the BECOME component is within the lower VP, so the VP has to be promoted to Spec, VoiceP.

(28)

\[
\text{AspP}
\vdash \text{ASST-T} \quad \text{Asp} \quad \text{BecP} \\
\quad \text{Asp} \quad \text{BecP}
\]

\[
\text{VoiceP}
\]

\[
\text{EVT-T} \quad \text{Bec} \quad 
\]

This assumption can also explain why some stative verbs such as ‘fear’ are allowed in the be-passive but not in the bei passive, as shown in (29-30).

(29) a. John was frightened by Bill.
    b. Bill was feared by John.

(30) a. Zhangsan bei Lisi xia-le yi-tiao.
    ‘Zhangsan got frightened by Lisi.’

b. *Lisi bei Zhangsan pa-le henjiu
    ‘Lisi got feared by Zhangsan for a long time.’

The verb frighten in (29a) and its Chinese counterpart xia in (30a) are causatives which denote a heterogeneous event that satisfy the requirements of BECOME components in these passives. As for (29b), G&G propose that the fear-type stative verbs can undergo verbal passivization because they have the possibility of being interpreted as resultant state by means of type shift from states to achievements when they combine with the BECOME component in the lower VP shell in (26), thus meeting the requirement of the BECOME component in a passive construction.

However, as I assumed, in Chinese passives, the BECOME component does not exist in the lower VP because a BECOME component is with the BecP above VoiceP but below AspP, and thus there is no need to add such a semantic element to VP for its promotion to Spec, VoiceP to anchor the event time. In addition, there cannot be two BECOME components in the structure since we cannot get the meaning of ‘become a becoming’.

Why is (30b) ungrammatical? Why cannot ‘fear’ verbs appear in the bei passive? I assume that immediately after this type of verbs enter into the computation, the BECOME component has to be added to coerce the result state meaning of the verbs. This can be done in the be passive in which the BECOME component is within the lower VP. However, in the bei passive, since the BECOME component is above VoiceP, it is too late for type shift to work.
Remaining as stative verbs, *fear*-type verbs are not compatible with the BECOME component in BecP above VoiceP, which require accomplishment predicates. Hence, stative verbs such as ‘fear’ cannot enter into the formation of Chinese *bei* passives.

5. Conclusions

In this paper, I studied the word order differences between the Chinese (short and local long) *bei* passives and English *be* passives based on the smuggling approach (Collins 2005) and a similar VP-movement approach (G&G 2009). The conclusions are drawn as follows:

First, the local long and short Chinese *bei* passives have the same underlying structure with the English *be* passive proposed by Collins (2005). However, different from English, Chinese does not involve smuggling.

Secondly, under the raising analysis of Chinese short or local long *bei* passives, since Chinese allows an object to be preposed to a position above VP but below TP, the VP can have an extra Spec to serve as the intermediate landing site for the internal argument to take, so its promotion does not really cross the external argument and hence no economy condition is violated. English does not have such kind of object preposing, no extra Spec, VP is available. Therefore, smuggling is adopted by the English *be*-passive as a last resort.

Thirdly, under the control analysis of Chinese short or local long *bei* passives, there is no movement of an overt object beyond *bei*. No island or minimality issues arise, so smuggling is not needed, hence disallowed.

Fourth, under G&G’s account, the BECOME component is as a decomposed component of *bei* which is above VoiceP. It can be associated with the temporal information directly without moving VP to Spec, VoiceP to satisfy the requirement of the event structure. However, in English, the BECOME element is within the lower VP which has to move to Spec, VoiceP to anchor the event time. This can also help explain the distribution of ‘fear’ verbs in English *be* passives and Chinese *bei* passives.

These circumstances then allow the Chinese passives to maintain ‘Kaynean word order par excellence’ as opposed to English (Huang 2008).

References


A CYCLIC LINEARIZATION APPROACH TO POLYNESIAN VP-REMNANT MOVEMENT

David J. Medeiros
University of Michigan

1. Introduction

This paper proposes a formal analysis of word order patterns in Polynesian, focusing especially on Hawaiian and Niuenean. I discuss Massam's (2001) influential VP-remnant movement account of verb-initial word order in the ergative/absolutive Niuenean, according to which VP-remnant formation is driven by case. I present several empirical and conceptual problems with Massam's analysis and analyze data from Hawaiian, which patterns with Niuenean in many respects, despite having a nom/acc case system. By providing a principled explanation for VP-remnant formation, I maintain a VP-remnant movement analysis for both languages. The primary theoretical contribution of this paper therefore concerns not VP-remnant movement, but VP-remnant formation. I argue that the process of VP-remnant formation is similar in some respects to Scandinavian object shift.

The cyclic linearization algorithm I develop adopts aspects of Fox and Pesetsky's (2005) system and also Müller's (2007) relativized and phase-based adaptation. I depart in some respects from Müller's relativization algorithm and argue that the syntactic representation is spelled-out at each transformational rule application (see e.g. Epstein and Seeley 2002), yielding a highly derivational model of linearization.

This paper is structured as follows. In the following section, I present basic word order facts from Niuenean and Hawaiian, establishing that both of these languages are predicate initial, although both VSO and VOS words orders are allowed. In section 3, I discuss Massam's (2001) VP-remnant movement account of Niuenean; this is followed by defending an incorporation account (contra Massam) for VOS word order. I argue that while Massam's VP-remnant movement account should be extended to Hawaiian, her analysis of VP-remnant formation, driven by case, faces empirical and theoretical problems. I propose a relativized cyclic linearization algorithm that forces VP-remnant formation in the derivation of VSO word order, divorcing VP-remnant formation from case.

2. Niuenean and Hawaiian are Predicate Initial Languages

Massam (2000, 2001) argues that Niuenean (1) is best understood as a predicate-initial language, based on data from non-verbal predicates (i.e. null copula, adjective, nominal, and...
prepositional predicates). Hawaiian (2) patterns with Niuean in this respect for these predicate types.

(1) a. Ko Mele e faiaoga. Niuean
    PRED Mele ABS teacher
    "The teacher is Mele."¹

    b. Ha he fale gagao a ia.
    PRED in house sick ABS she
    "She is in the hospital." (Massam 2000)

(2) a. He kumu kula 'o Noelani. Hawaiian
    a teacher-school SUBJ Noelani
    "Noelani is a teacher."

    b. Ua ha'i 'o Kekoa he kumu kula 'o Noelani.
    PERF say SUBJ Kekoa a teacher-school SUBJ Noelani
    "Kekoa said that Noelani is a teacher."

    c. Hau'oli 'o Kekoa.
    happy SUBJ Kekoa
    "Kekoa is happy." (author field notes, FN)

Although unexpected from the perspective of a predicate initial analysis, the word order typically discussed in descriptions of both languages is VSO, when the main verb is lexical (3,4).

(3) Ua ku'ai 'o Kekoa i ka ia. Hawaiian
    PERF buy SUBJ Kekoa OBJ the fish
    "Kekoa bought a fish."

(4) Takafaga tumau ni e ia e tau ika. Niuean
    hunt always EMPH ERG he ABS PL fish
    "He is always fishing." (Massam 2001, Siter 1980)

However, VOS is also allowed when the object is non-specific (5a). Case markers and articles are excluded with VOS (5a-c). VOS is also allowed in embedded clauses (5d).

(5) a. Inu *(ana 'o) Noelani i ke kope hu'ihu'i. Hawaiian
    drink DIR SUBJ Noelani OBJ the coffee cold
    "Noelani is drinking cold coffee."

    b. Inu kope hu'ihu'i 'o Noelani. drink coffee cold SUBJ Noelani
    "Noelani is drinking cold coffee."

c. *Inu i ke kope hu'ihu'i 'o Noelani.
   drink OBJ the coffee cold SUBJ Noelani
   "Noelani is drinking cold coffee."

d. No'ono'o 'o Kekoa ke 'ai poi nei 'o Noelani.
   think SUBJ Kekoa PRES eat poi DIR SUBJ Noelani
   "Kekoa thinks that Noelani is eating poi." (FN)

In Hawaiian, adverbs like mau 'always', which normally intervene between the verb and subject (6a), follow the V-NP sequence in V-NP_{obj}-S word order (6b), suggesting that the entire predicate has fronted as a single constituent in V-NP_{obj}-S word order.

(6) a. Holoholo mau 'o Kehau ma ke kaona. Hawaiian
   cruise always SUBJ Kehau in the town
   "Kehau always cruises in town." (Cleeland, 1994)

   b. Inu (*i ke) kope mau 'o Noelani.
      drink coffee always SUBJ Noelani
      "Noelani always drinks coffee."

   c. *Inu mau kope 'o Noelani.
      drink always coffee SUBJ Noelani
      "Noelani always drinks coffee." (FN)

While Hawaiian is a nom/acc language, Niuean has erg/abs case marking, and absolutive case appears on the agent in the V-NP_{obj}-S word order (i.e. case marking behaves intransitively for Niuean V-NP_{obj}-S). Niuean is syntactically similar to Hawaiian with respect to adverbials.

(7) Takafaga ika tumau ni a ia. Niuean
   hunt fish always EMPH ABS he
   "He is always fishing." (Massam, 2001)

For both languages, tense/aspect marking is accomplished by pre- and post-verbal particles. For V-NP_{obj}-S order, TAM markers cannot intervene between V and O.

(8) E inu kope hu'ihu'i (nei) 'o Noelani. Hawaiian
    IMPERF drink coffee cold (DIR) SUBJ Noelani
    "Noelani is drinking cold coffee." (FN)

(9) Ne inu kofe kono a Mele. Niuean
    PST drink coffee bitter ABS Mele
    "Mary drank bitter coffee." (Massam, 2001)

On an empirical level, these data show that descriptions like VSO and VOS can be misleading. Rather, V-NP_{obj}-S and V-S-DP_{obj} are better (if still superficial) characterizations of these languages.

More importantly for the current paper, the data presented above show that Hawaiian is similar in many respects to the better-studied Niuean. In particular, the word order facts, including the distribution of adverbs and TAM markers in both VSO and VOS constructions, suggests that whatever property accounts for V-NP_{obj}-S and V-S-DP_{obj} word orders in Niuean
should also account for the same word order patterns in Hawaiian. In the next section, I present Massam's (2001) influential analysis of Niuean word order, which, I argue, can explain many of the relevant facts for Hawaiian as well. However, Massam's analysis is crucially tied to case properties, arguably the primary locus of morpho-syntactic variation between the two languages. In order to develop a unified analysis of the two languages, I rework Massam's system in section 4 in an analysis that does not make reference to case.

3. VP-remnant Movement and Subject/Non-Subject Asymmetry


Massam (2001) unifies the predicate initial property of Niuean by proposing VP-movement to Spec, TP, the EPP feature of which she specifies as [PREDICATE], not [D]. While I will maintain Massam's VP-(remnant) movement hypothesis, I will reject her analysis of VP-remnant formation. In order to derive V-S-DP<sub>obj</sub>, Massam proposes VP-remnant movement. According to Massam, everything but V must vacate VP prior to VP-movement to Spec, IP/TP. Massam ties VP-remnant formation to case; when a DP (not NP) object is generated, it must raise to Spec, AbsP (10).

(10) Niuean Transitive VSO clause (Massam 2001)

Massam (2010) updates this model, including vP. The details of the remnant formation are similar, with vP raising to Spec, TP (11)

(11) [CP C [TP [vP v+V [vP tv t_{i}]] T [ErgP S Erg^{o} [AbsP KP_{i} Abs^{o} t_{vP}]]]] (Massam, 2010)

Under Massam's analysis, in V-NP<sub>obj</sub>-S clauses the complement of the verb is not a DP, but rather an NP. Therefore, there is no motivation for the verbal complement to move for case reasons. If an NP is generated, then the entire VP moves to Spec, TP (12).

Rackowski and Travis (2000) and Massam (2000) also offer VP-remnant accounts of Niuean.
Massam assumes that "agents can be variously generated in Spec of ErgP or Spec of AbsP ... an argument is generated in $v^{\text{Max}}$ (hence ergative) only if (a) it is an agent and (b) absolutive case has been checked" (Massam 2001, footnote 15). Therefore, there is an unexplained co-occurrence pattern between the complement of V and the presence of ErgP (or at least ergative marking on v); whenever an NP is generated as the complement of the verb, an ErgP head must not be generated, as no element would be present in the derivation to check the ergative case feature.

Massam terms the V-NP$_{\text{obj}}$S construction 'Pseudo-Noun-Incorporation (PNI),' since the V-NP$_{\text{obj}}$ sequence has several properties characteristic of incorporation. However, to be clear, PNI is not incorporation for Massam; rather this is VP-movement just in those instances when an NP is generated as the sister of V.

Massam's VP-remnant movement hypothesis has been very influential in Polynesian linguistics, since it unifies the predicate initial data seen above. Nevertheless, two of Massam's central claims, (i) that incorporation does not apply, and especially (ii) that VP-remnant formation is driven by case, have been challenged, calling into question the overall analysis. To my knowledge, no formal solution for (ii) has been developed, and this will be my goal in section 4.

### 3.2. Incorporation in V-NP$_{\text{obj}}$S Word Order

Massam's arguments against incorporation largely follow from theoretical constraints on X-bar theory, many of which have been challenged by Minimalist research (e.g. Chomsky 1995 and much subsequent work). Most importantly, Massam claims that "a head is distinct from a phrase ... and a head cannot contain a phrase." Chung and Ladusaw (2004) take a different approach with respect to V-NP$_{\text{obj}}$S; they discuss the Niuean data and also Maori, which patterns with Niuean in allowing full NPs in the possible incorporation structure (13).

(13) E rukuruku [koura nunui] ana ia.     Maori
    TAM dive crayfish big DIR she
"She is diving for big crayfish." (Bauer 1997)

Because "nothing, not even pro forms" such as ia or tense/directional particles such as ana can separate the verb from the NP object, they argue that Massam's evidence "is consistent with an incorporation analysis...consequently, we believe that Niuean does have incorporation" (Chung and Ladusaw, 2004, original emphasis).

In light of Chung and Ladusaw (2004), I suggest that the argument for 'real' V+NP incorporation in languages such as Niuean, Maori, and Hawaiian is at least as strong as those against incorporation. Putting aside further details regarding how this incorporation takes
place (see again Chung and Ladusaw (2004) and Massam (2009) for discussion), I suggest that PNI structures in these languages are bona-fide instances of incorporation (i.e. forming a complex \( V^0 \)), possibly as represented in (14).

\[(14) \text{(Chung 1988)}\]

Furthermore, a structure such as this provides the kind of representation that is justified by the linearization-based account of VP-remnant formation developed below. I will assume that is because NP is dominated by an \( X^0 \) element, i.e. 'below' the word level under incorporation and therefore invisible to the linearization process which feeds PF from the syntactic representation.

### 3.3. The V-NP<sub>obj</sub>-S/V-S-DP<sub>obj</sub> and Subject/Non-Subject Asymmetry

If V-NP<sub>obj</sub>-S involves incorporation, then this is consistent with Chung's (1998) claim (not formalized) that subjects generally precede non-subjects in Maori. For example, Maori (15) and Tongan (16), like Hawaiian and Niuean, also have the V-NP<sub>obj</sub>-S/V-S-DP<sub>obj</sub> alternation, suggesting that at least this aspect of the grammar is part of a larger trend within Polynesia.

\[(15)\]

a. Kei te a Rewi i nga poaka
   IMPERF PERS Rewi OBJ PL pig
   "Rewi is feeding the pigs."

b. E karanga manuhiri ana ia.
   PRES call visitor DIR she
   "She is welcoming visitors." (Chung and Ladusaw, 2004, citing Waititi 1962 and Bauer 1997)

\[(16)\]

a. Na'e ma'u 'e Sione 'a e ika.
   PAST get ERG Sione ABS the fish
   "Sione got the fish."

b. Na'e inu kava malohi 'a e kau siana.
   PAST drink kava strong ABS the PL man
   "The men drank alcohol." (Otsuka 2005)

While there is some debate as to whether V-NP<sub>obj</sub>-S order is as productive in Tongan as compared to Niuean (Otsuka 2005, Ball 2005), the facts regarding functional marking of the object are the same.

Various additional constructions also support a general subject/non-subject asymmetry, such that subjects precede non-subjects. For example, several researchers have argued that

(17) *I na pua mae e waele ai ke ali'i i ka paka Hawaiian
OBJ PL flower wilt PRES weed RESPRO the chief PREP the park
"The wilted flowers the chief will weed in the park" (Hawkins, 1982)

When an object appears sentence initially, it is part of a clear pseudo-cleft; compare VSO (18a) with (18b), in which the wh-object appears in a pseudo-cleft construction.

(18) a. Ua ku'ai 'o Kekoa i ka i'a.
PERF buy SUBJ Kekoa OBJ the fish.
"Kekoa bought a fish."

b. He aka ka mea a Kekoa i ku'ai ai?
A what the thing POSS Kekoa PERF buy RESPRO
"What is the thing that Kekoa bought?" (FN)

While objects can only appear initially in bi-clausal structures (i.e. pseudo-clefts), subjects can appear sentence-initially in a number of constructions which are not obviously bi-clausal. Some of these have been argued to be clefts, involve raising, or be A-bar movement in various Polynesian languages. While the following constructions may involve clefts (and this is arguably the consensus position, see also Potsdam 2009), the overt evidence is far less clear as compared to the object-fronting examples. First, in Hawaiian, the subject in subject wh-questions can appear sentence initially (19) without clear evidence of a cleft/pseudo-cleft (compare to (18b)).

(19) 'O wai i ku'ai i ka i'a.
SUBJ who PAST buy OBJ the fish
"Who bought a fish?" (FN)

Second, Hawaiian subject pronouns must precede the lexical verb under negation. (20a) shows the baseline sentence and (20b) the sentence including negation.

(20) a. Ua hele 'o-ia
PERF go SUBJ-he
"He has gone."

b. A'ole 'o-ia i hele.
not SUBJ-he PERF go
"He didn't go." (Elbert & Pukui 1979)

Finally, the 'actor emphatic' is another example of subject fronting that is unavailable for objects. According to Hawkins (1979), the fronted element is both semantic agent and syntactic subject.

(21) a. Ua heluhelu ke keiki i ka puke.
PERF read the child OBJ the book
"The child read the book."
b. Na ke keiki i heluhelu i ka puke.
NA the child PERF read OBJ the book
"THE CHILD read the book." (Hawkins 1979)

Independently of whether (19-21) involve cleft structures, the claim that subjects precede objects holds for derivations involving movement. The only superficial exceptions are the cases of object clefts and V-NP incorporation, which do not involve movement of the object over the subject, since the incorporation structure is understood as a complex word and the object pseudo-cleft is likely base-generated. If (19-21) are not clefts, then a subject/non-subject asymmetry holds for an even wider range of constructions. To the extent that a subject/non-subject asymmetry is consistent with the V-NP$_{obj}$-S/V-S-DP$_{obj}$ alternation discussed above, a theory of the latter should also explain the former, and the analysis developed in section 4 below can account for both patterns of data.

3.4. Empirical and Conceptual Problems with the VP-remnant Movement Hypothesis

Returning now to Massam's (2001) VP-remnant movement account of Niuean, in this section I present some empirical and conceptual problems with Massam's account. While Massam's claim that VP moves to Spec, TP is attractive, her analysis has been criticized; much of this criticism, however, does not involve VP-remnant movement per se, but rather the formation of the VP-remnant prior to VP-remnant movement. For instance, a well-known (Chung 2005, McCloskey 2005 and others) empirical issue involves sentences in which the verb takes a CP complement. This yields V-S-CP order (22-24), suggesting that CP must vacate VP along the same lines as DP. However, there is no evidence that CP needs case in the relevant languages.

(22) Ne manatu e Mataginifale [ko e mena fai mata-fohi haku hiapo]... Niuean
PST think ABS Mataginifale PRED ABS thing have blade-scraper scratch tapa-plant
"Mataginfale remembered that she had the blade of the tapa plant scraper...
(Niue: 1982, cited by Massam 2001)

(23) Ua no’ono’o ‘o Kekoa [ke ‘ai nei ‘o Noelani i ka poi]. Hawaiian
PERF think SUBJ Kekoa PRES eat DIR SUBJ Noelani OBJ the poi
"Kekoa thought that Noelani is eating poi."

(24) E ninua maila ‘o Kekoa [ina ‘olelo Hawai’i ‘o Noelani].
IMPERF ask DIR SUBJ Kekoa if speak Hawaiian SUBJ Noelani
"Kekoa is asking if Noelani speaks Hawaiian." (FN)

Hawaiian modal (25) and impersonal verbs (26) display a similar pattern.

(25) Pono i ke keiki [ke ha’i mai i ka mo’olelo].
should PREP the child PRES tell DIR OBJ the story
"The child should tell the story." [It is necessary for the child that (he) tell the story.]
(Hawkins, 1979)

(26) Maopopo ia Noelani [e hele mai ana ’o Kekoa].
known to Noelani IMPERF go DIR DIR SUBJ Kekoa
"It is clear to Noelani that Kekoa is coming." (FN)
Double objects (27) and (possible) PP complements (28) also appear after subjects in Hawaiian.

(27) Ua lawe aku 'o Kaipo i ka-na wahine i ka hale ola.
PERF take DIR SUBJ Kaipo OBJ his wife PREP the hospital
"Kaipo took his wife to the hospital." (Hawkins, 1979)

(28) Ke ha'awi aku nei au i keia ipu ia 'oe.
PRES give DIR DIR I OBJ this gourd IN.OBJ you
"I am giving this gourd to you." (Hawkins, 1979)

If CP (22-24) does not need case, then there is no motivation for CP to vacate VP under Massam's analysis, nor is it clear how to extend her analysis to cover (25-28).

There are conceptual problems as well. First, as noted above, there is an unexplained co-occurrence pattern in Massam's system; whenever an NP is a sister to V, ErgP must not be generated. Also, under Chomsky's (2000, 2008) model of long-distance Agree, case and phi-features do not motivate movement. Finally, Hawaiian (and Maori) lack the erg/abs case marking that is at the heart of Massam's analysis. An ad-hoc solution to this problem would be to suggest that object DPs do not get case valued in-situ, but must raise to the specifier of a functional head such as AgrO, although there is no independent evidence for such a projection in Hawaiian. Instead of abandoning the VP-remnant movement hypothesis, I develop a model of VP-remnant formation which does not involve case, allowing a unified analysis for Niuean and Hawaiian.

4. A Relativized Cyclic Linearization Approach

If VP-remnant formation in Polynesian is not case-driven, this movement may be analogous (under at least some analyses) to a much better-studied phenomena, namely Scandinavian object shift. In particular, several researchers have taken a 'shape conservation' approach to Holmberg's Generalization, i.e. that an object of a verb may shift only if that verb undergoes raising. The idea here is that object shift preserves the shape of the predicate only if the verb raises, where 'shape' is often understood to be a property of linear order of some domain.

(29) a. Jag kysste henne inte [vp tv t0].
   I kissed her not
   "I didn't kiss her."

   b. *Jag har henne inte [vp kysst t0].
   I have her not kissed
   "I have not kissed her."

While the object movement in the grammatical (29a) is optional, this may be analogous to the obligatory movement of verbal complements in Polynesian, since (by hypothesis) neither movement is driven by case.

The key for the Hawaiian data will be to develop a model in which subjects and objects maintain their base linear order (assuming underlying SVO), while the verb may raise past the subject; intuitively, V-XP_{comp}-S reverses the underlying word order while V-S-XP_{comp} preserves it. Under such an analysis, verbal complements can only raise past subjects if incorporation has taken place, such that (only) NP complements to V escape the shape conservation effect under incorporation. The proposal I develop here follows the tradition of
Fox and Pesetsky (2005) and Müller (2007), both of whom formalize the shape conservation intuition in terms of cyclic linearization.

Very briefly, Fox and Pesetsky (2005) account for a wide range of movement restrictions in a model in which the syntactic representation undergoes spell-out at CP and VP in Germanic (VP is parameterized with vP in other languages). Ungrammaticality occurs when contradictory statements are generated at different spell-out domains (although these can be deleted, allowing e.g. island amelioration). Under Fox and Pesetsky's system, Holmberg's Generalization is derived; V and O are strictly ordered within VP, such that O may shift only if V also raises.

Müller (2007) claims that Fox and Pesetsky's system is "both too strong ... and too weak," as this makes wrong predictions for a number of constructions while at the same time failing to derive shape conservation effects between subjects and objects in Germanic, which are arguably as strong or stronger than those between objects and verbs (see also responses to Fox andPesetsky article in *Theoretical Linguistics 31*). Müller therefore proposes that spell-out nodes correspond to Chomsky's phasal nodes CP/vP (strengthening the system by allowing subjects to enter ordering relationships with elements in VP), and also that linearization relationships should be relativized with respect to Merge Status:

(30) **Merge status (Müller 2007), to be amended**

a. A category $\gamma$ in a position $P$ has Merge status $[-\psi]$ iff (i) or (ii):

(i) $\gamma$ is merged in $P$, and $\gamma$ is required in $P$ by a non-local Merge inducing feature.

(ii) $\gamma$ is dominated by (a segment of) a category with Merge status $[-\psi]$.

b. A category $\gamma$ in a position $P$ has Merge status $[+\psi]$ iff (i) or (ii):

(i) $\gamma$ is merged in $P$, and $\gamma$ is not required in $P$ by a non-local Merge inducing feature.

(ii) $\gamma$ is not dominated by (a segment of) a category with Merge status $[-\psi]$.

Merge inducing features include subcategorization features, EPP on T, semantic operators (e.g. wh-feature), etc. In essence, an item is $[-\psi]$ if it is in an intermediate landing site. In this system, elements with like Merge Status are relativized with respect to each other; $[+\psi]$ elements create orderings with like elements and likewise for $[-\psi]$ elements. Look-ahead is avoided via Phase Balance, which requires derivational access to the lexicon.

I agree with Müller that Fox and Pesetsky's system is both too strong and too weak; however, I reject stipulated spell-out domains like Phases and instead suggest that linearization (and not Phase-based spell-out) occurs with each transformational rule application. I also offer two simplifications of Müller's system. First, I amend Merge Status as follows.

(31) **Merge Status, final version**

A terminal node $\gamma$ in a position $P$ has Merge Status $[-\psi]$ if it is required in $P$ by a non-local Merge inducing feature; otherwise it has Merge Status $[+\psi]$.

Additionally, I propose that only elements with $[+\psi]$ status are visible to the linearization process.

(32) **Relativized Linearization**

Linearization generates for categories $x$, $y$ an ordering statement $<x,y>$ iff $x$ and $y$ have $[+\psi]$ Merge status.
4.1. Deriving Obligatory XP-shift in Polynesian

A crucial aspect of the relativization scheme is the notion 'Merge inducing' feature (31). Let's consider how this concept applies to various elements in different positions for Hawaiian and Niuean, under the assumption that phi- and case-features do not cause movement under Agree.

(33) Merge Status of terminal elements by position in Hawaiian/Niuean
a. XP sister to V: [+ψ], it is required locally by the subcategorization feature of V
b. 'object shifted' XP: [+ψ], it is not required by a non-local feature (i.e. not an 'intermediate' site)
c. V in base position: [-ψ], it is not required by any local syntactic feature
d. v in base position: ø, because v° is null
e. v+V in v°: [-ψ], it is not required by any local syntactic feature
f. v+V in Spec, TP: [+ψ], it is required by EPP

g. S in base position: [+ψ], it does not raise, therefore not required by a non-local feature

Assuming that linear order is related to asymmetric c-command (Kayne, 1994), and that T is null (arguably a set of affixes that undergoes T-C movement, Massam (2010)), the formal burden for this analysis is to show that an ordering contradiction is generated whenever a DP/CP complement of V does not undergo object shift, blocking e.g. the following example (repeated from above).

(34) *Inu i ke kope hu'ihu'i 'o Noelani.
    drink OBJ the   coffee cold   SUBJ Noelani
    "Noelani is drinking cold coffee."  (FN)

First consider a derivation, like that proposed in Massam (2001), in which VP moves to Spec, TP. Crucially, the subject is [+ψ] in Spec, vP (where it enters the structure), entering an ordering relationship immediately with the [+ψ] XP sister of V (35c). In addition to forcing XP movement to a position outside VP (but lower than Spec, vP), this also derives the subject/non-subject asymmetry discussed in section 3.3 above.

(35) Derivation of *[TP [vP V DP/XP] T [vP S ...]]: VP moves to Spec, TP
a. [vP V XP] → ∅ (read 'no ordering,' here due to lack of asymmetry and the presence of only one [+ψ] item in the derivation)

b. [vP v [vP V XP]] → ∅ (v° is null and therefore not marked for ψ-status)
   [-ψ] [+ψ]

c. [vP S v [vP V XP]] → S < XP (S and XP share [+ψ] status; S c-commands XP)
   [+ψ] [-ψ] [+ψ] this derives the subject/non-subject asymmetry noted above

d. [TP T [vP S v [vP V XP]]] → ∅
   [+ψ] [-ψ] [+ψ]

e. [TP [vP V XP] T [vP S v [tVP]]] → *[XP < S (compare to (c); VP asymmetrically c-commands S, forcing XP<S order)]
f. *[CP C [TP [VP V XP] T [vP S \(v\) t\(V\)]]]] \(\rightarrow\) C < V
\[ [+\psi] \quad [+\psi] \quad [+\psi] \quad [+\psi] \]

This derivation introduces an ordering contradiction, correctly predicting ungrammaticality. A VP-fronting derivation can then only be grammatical if either XP=NP, such that incorporation occurs, or if XP moves to some position external to VP but lower than S.

Another possible derivation, in which vP moves to Spec, TP, is similarly blocked.

(36) Derivation of *[TP [vP v+V [VP t\(V\) DP/XP]] T [vP S ...]]: minimal vP dominating v moves to Spec, TP

a. [vP V XP] \(\rightarrow\) \(\emptyset\)
\[ [-\psi] \quad [+\psi] \]

b. [vP v+V [vP t\(V\) XP]] \(\rightarrow\) \(\emptyset\)
\[ [-\psi] \quad [+\psi] \]

c. [vP S v+V [vP t\(V\) XP]] \(\rightarrow\) S < XP
\[ [+\psi] \quad [-\psi] \quad [+\psi] \]

d. [TP T [vP S v+V [vP t\(V\) XP]]]] \(\rightarrow\) \(\emptyset\)
\[ [+\psi] \quad [-\psi] \quad [+\psi] \]

e. [TP [vP v+V [vP t\(V\) XP]] T [vP S t\(vP\)]]] \(\rightarrow\) *XP < S
\[ [+\psi] \quad [+\psi] \quad [+\psi] \quad [+\psi] \]

f. *[CP C [TP [vP v+V [vP t\(V\) XP]] T [vP S t\(vP\)]]]] \(\rightarrow\) C < V
\[ [+\psi] \quad [+\psi] \quad [+\psi] \quad [+\psi] \quad [+\psi] \quad [+\psi] \quad [+\psi] \]

Given the linearization model, the structure of a V-S-DP\(_{obj}\)/V-S-CP may be broadly as follows.

(37) [CP C [TP [vP v+V [vP t\(V\) t\(XP\)]] T [vP S XP t\(vP\)]]]]

\(\text{Inu} \quad \text{\textquotesingle o Noelani} \quad \text{i ke kope...} \)
Drinks Noelani the coffee

4.2. Theoretical Advantages of the Proposed System

Given that the most concrete prior proposal in this domain (Massam, 2001) tied VP-remnant formation to case, and also considering that case properties are the primary locus of syntactic variation between Niuean and Hawaiian, this proposal offers a unified analysis of Niuean and Hawaiian (and, I argue, other Polynesian languages such as Maori and Tongan) with respect to VP-remnant formation.

While the proposed system allows movement for case in Niuean as proposed by Massam, it does not require it. Therefore it is possible to eliminate null case heads from the representation of Niuean syntax, while allowing alternative analyses for Niuean case marking (see Legate 2008 for an alternative proposal regarding Niuean case marking). Moreover, DP and CP movement out of VP is accounted for in the same way. In both cases, the inability of VP internal elements to escape VP is derived, due to the [+\(\psi\)] property of the subject when in-situ in Spec, vP.

- 152 -
4.3. Extending the Model to Scandinavian Object Shift

In order to extend the model to Scandinavian object shift, consider the $\psi$-status of the subject within the verbal domain. Syntactic and semantic evidence suggests that subjects stay in-situ in at least Hawaiian (there is no evidence of raising and subjects are allowed to be indefinite and non-specific). In Germanic, however, one can argue that the subject Merges into Spec, vP primarily in order to satisfy T’s EPP feature; i.e. Spec, vP can be (under some analyses) considered a kind of 'intermediate position' for subjects in these languages. If this can be substantiated further, then subjects may be taken to be [-$\psi$] in their base position in Germanic. Under this account, subjects do enter into ordering relationships with verbs and objects (contra Fox and Pesetsky), but only at a later stage of the derivation (i.e. at Spec, TP) when compared to Polynesian.

Returning to Holmberg’s Generalization, the possibility of object shift is straightforwardly derived within the current proposal. (38) shows the relevant derivation.

(38) Grammatical instance of object shift

a. $[\text{vp } V \quad O] \rightarrow \emptyset$  
   [-$\psi$]  [+\$\psi$]

b. $[\text{vp } v+V [\text{vp } tv \quad O]] \rightarrow \emptyset$  
   [-$\psi$]  [+\$\psi$]  \text{(v+V will undergo further raising, suggesting it is required in this position by a non-local feature)}

c. $[\text{vp } \text{NEG } v+V [\text{vp } tv \quad O]] \rightarrow \emptyset$  
   [-$\psi$]  [+\$\psi$]

d. $[\text{vp } S [\text{vp } \text{NEG } v+V [\text{vp } tv \quad O]]] \rightarrow \emptyset$  \text{(S and v+V do not order due to [-$\psi$] status)}  
   [-$\psi$]  [-$\psi$]  [+\$\psi$]

At this point the object shifts to a position higher than VP (and Neg) but lower than TP. This can be achieved within vP either by 'tucking in,' as shown here or, just as easily, by first moving O before Merge of the subject (somewhat reversing the steps in this derivation).

(38) e. $[\text{vp } S [\text{vp } O [\text{vp } \text{NEG } v+V [\text{vp } tv \quad t_o]]]] \rightarrow \emptyset$  \text{(O maintains [+\$\psi$] status, as this is not an intermediate landing site)}  
   [-$\psi$]  [+\$\psi$]  [-$\psi$]

f. $[\text{tp } \text{T}[v+V] [\text{tp } t_s [\text{vp } O [\text{vp } \text{NEG } t_s+v [\text{vp } tv \quad t_o]]]]] \rightarrow S<v+V, v+V<O$  
   [+\$\psi$]  [+\$\psi$]  [+\$\psi$]

Note that this model derives grammatical object shift in a rather different way as compared to Fox and Pesetsky. Whereas in their model verb raising after object shift reconstructs the original V<O order, in this model object shift is allowed because V and O never enter an ordering relationship within VP in the first place, due to relativization.

For the ungrammatical case, the lexical verb does not raise, and therefore V will be marked [+\$\psi$] within vP. (The alternate $\psi$-status of V within vP in the two derivations (38) and (39) could be deduced given a feature-based understanding of head movement, such as the one proposed in Matushansky’s (2006), in which head movement is triggered by Agree; examination of the feature content of the relevant heads may then correctly yield the necessary $\psi$-status values). This yields a strict V<O ordering early in the derivation.
(39) Relevant steps of an ungrammatical instance of object shift derivation

a. \([vp \ V \ O] \rightarrow \emptyset\)

\([-\psi] \ [+\psi]\]

b. \([vp \ v+V [vp \ t_v \ O]] \rightarrow V < O\)

\[+\psi] \ [+\psi]\]

Due to the nature of the cyclic linearization, (39b) establishes V< O order that cannot be reversed via object shift. Therefore, object shift is correctly barred when the lexical verb does not raise.

5. Conclusions

I argued that the V-NP_{obj}-S/V-S-DP_{obj} alternation in Polynesian can be understood as a type of subject/non-subject asymmetry, when V-NP_{obj}-S is understood as bona-fide incorporation. This is especially true when CP complements of V are considered; for both DP (V-S-DP_{obj}) and CP (V-S-CP), the subject must precede the VP internal material. While VP-remnant movement along the lines of Massam (2001) offers an attractive account of predicate initial structures in several Polynesian languages, VP-remnant formation as understood by Massam's case-based account introduces several empirical and conceptual problems; CPs as well as DPs must vacate VP, and the word order facts are very similar for both nom/acc and erg/abs languages.

I therefore proposed a relativized cyclic linearization algorithm that offers a syntax-phonology interface solution for the problem of VP-remnant formation; this account does not depend on case properties. This model may extend to further subject/object asymmetries in the relevant languages. The relativization algorithm can accommodate cross-linguistic variation allowing a possible extension of the model to Scandinavian object shift.

References


THE APPARENT LACK OF ADJECTIVAL CATEGORY IN MALAYALAM AND OTHER RELATED LANGUAGES

Mythili Menon
University of Southern California

1. Introduction

Over the years, one of the tasks of generative theory has been to find and explain language universals. One such universal is the notion of primitive lexical categories, namely Noun (N), Verb (V), Adjective (A) and Preposition (P) (Chomsky 1970). Every language has words belonging to these four classes, and they are listed as such and categorized in the lexicon. In early generative grammar, these four categories were characterized in terms of binary distinctions of N and V features. An adjective in this view is [+N, +V]. Currently the most prominent contender for the Universalist approach to lexical categories is Baker (2003, 2005) who argues that at least N, V, and A, are universal, although A can be quite varied in realization.

Recently the Universalist view has been challenged from data pertaining to variations found cross-linguistically across A and P. Particularly, in the Dravidian literature the question of whether the language family indeed has a separate lexical category of adjectives has remained controversial (see e.g. Zvelebil (1990: 27)). More recently from a functional perspective, Bhat (1994) has argued that Dravidian does lexicalize the adjectival category. Amritavalli and Jayaseelan (2003), Jayaseelan (2007) argue in a Lexical Relational Structure (LRS) approach (Hale and Keyser 1993) for an incorporation account of adjectives. For them universally, adjectives are created with the incorporation of a noun into a preposition or a Case head. Thus, the Dravidian literature is still divided amongst the view as to whether there is a separate lexical category for adjective.

This paper contributes to the discussion by arguing with data from Malayalam and other Dravidian languages that A cannot be universal since there is no independent class of adjectives in Dravidian. More specifically, adjectives are not found in the lexicon nor are they created in syntax. An adjectival-like construction can be syntactically created for the purpose of attributive modification and predication. With the help of verbal and nominal heads, a relativization structure is created for attributive modification, and a nominalization structure is created in the case of predication. The lexicon comprises only of roots.

Adjective-like constructions in Dravidian are built from roots that denote primitive property concepts of type $e^k$ (kinds). They thus need additional structure to be converted into what looks like adjectival constructions. Empirically, I show contra Baker (2003, 2005), Pearson (forthcoming) that A cannot be a universal category. A new theory for the syntax of the adjective is sketched, using tenets of Distributive morphology (Marantz 1997, Borer 2003) and the feature sharing system (Frampton and Gutmann 2006, Pesetsky and Torrego 2007). A

---

* I would like to thank Roumyana Pancheva, Hajime Hoji, Andrew Simpson, Audrey Li, Akira Watanabe, Mamoru Saito, the audience of Glow-in-Asia IX, and three anonymous abstract reviewers for valuable discussions and comments. Any remaining errors are solely mine.
1 Any use of the term “Adjective refers” to the lexical category of adjectives as found in English-type languages. I am not committed to whether the lexical A comes out of the lexicon as an A or whether it is a root combining with a lexically specified a head.
The desired consequence of the new system is the result obtained for case assignment and agreement morphology. It explains the dative/nominative case alternation in predicative constructions in Dravidian. Some of the consequences of the theory include the presence of only nominal and verbal comparatives, and the absence of resultative secondary predications in Dravidian.

The paper is organized as follows. In the next section I will look at the status of adjectives in Korean, Japanese, and some Bantu languages. In § 3.0, I will then introduce the basic paradigm concentrating on data from Malayalam. In § 4.0, the analysis will be sketched by crucially looking at the semantics and the syntactic representations. In § 5.0, I will address the residual issues and problems to be addressed for future research and conclude.

2. Missing Attributive Adjectives in Japanese and Other Languages

One of the canonical positions in which an adjective can appear is the attributive position. It is known that in some languages, attributive adjectives are missing. Japanese, Korean, Slave, Ika, and other Bantu languages belong to this category. In Japanese, which we turn to presently, the status of adjectives is as controversial as it is in the Dravidian literature. Korean attributive adjectives have been argued to be concealed relative clauses (Kim 2002). In Slave, adjectives appear predicatively but not in the attributive position unless relativized. In Ika, the use of a copular verbal element mediates the presence of the adjective in the attributive positions. Where then are the missing attributive adjectives?

2.1. Japanese Adjectives

The status of Japanese adjectives is controversial. There are two kinds of adjectives discussed in the literature. The first one is called the true adjective (Miyagawa 1987, Murasagi 1990) and in Chomsky’s (1970) terms is categorized as [+N, +V]. The other category comprises of the “verbal adjective” which is categorized as +V.

\[
\begin{array}{ll}
\text{True Adjective} & \text{Verbal Adjective} \\
\text{a. Kirei} & \text{b. Utsukushi}
\end{array}
\]

\[
\text{‘beautiful’} \quad \text{‘beautiful’}
\]

It is often assumed that the adjectives belonging to (1a) can appear as attributive modifiers without the help of any additional morphology. Verbal adjectives, on the other hand, can appear only with the help of a copula in the attributive position.

\[
\begin{array}{ll}
\text{uitsukushi-*(i) onna} & \text{Japanese} \\
\text{beautiful-PRES woman}
\end{array}
\]

Baker (2003) notes that the presence of the copular element makes (2) a relative clause structure and –i does not signal an attributive modifier. However, he wishes to maintain the view that (1b) type verbal adjectives in fact behave like attributive modifiers and are not similar to the characteristic functions of verbs. The diagnostics include resultative secondary predications (RSPs), the complement position of a degree word such as ‘too’, ‘as’, and unaccusativity predicates. In English, only adjectives can appear in RSPs. Nouns and verbs are unable to do so.

\[
\begin{array}{ll}
\text{a. I beat the metal flat} & \text{(AP)} \\
\text{b. * I beat the metal broke} & \text{(VP)} \\
\text{c. * I beat the metal (a) sword} & \text{(NP)}
\end{array}
\]
Utsukushi-type adjectives can appear in RSPs suggesting they are adjective-like in their behavior and unlike verbs in that sense (Ohkado 1991, Washio 1997).

(4) a. Taroo-ga kami-o mizika-ku kit-ta.
   Taro-NOM hair-ACC short-AFF cut-PST
   ‘Taro cut the hair short.’

 b. #Taroo-ga kami-o ochi(-te) kit-ta
   Taro-NOM hair-ACC fall-AFF cut-PST
   ‘Taro cut the hair so that it fell.’

In English, the complement position of dedicated degree words such as ‘too’ and ‘as’ necessarily has to be occupied by an adjective. Similarly, Utsukushi-type adjectives can appear in the complement position of a degree word suggesting their behavior is unlike that of verbs.

(5) a. Mary is too smart (to make such a mistake).

   b. * Mary is too (a) genius (to make such a mistake).

(6) a. Hanako-ga totemo utsukusi-i. (A)
   Hanako-NOM very beautiful-PRES.
   ‘Hanako is very beautiful.’

   b. * Hanako-ga totemo sensei-da. (N)
   Hanako-NOM very teacher-COP
   ‘Hanako is very (much a) teacher.’

   c. * Hanako-ga totemo okasi-o tabe-ru. (V)
   Hanako-NOM very sweets-ACC eat-PRES
   ‘Hanako very (much) eats sweets.’

A point however which Baker (2003) does not make is the fact that even (1a) needs the presence of a particle (in traditional grammar this could be a copula2) to obligatorily be present in order for the adjective to attributively modify a noun. Thus, for our purposes it is important to note that without the mediation of an extra particle, attributive modification is not possible for either type of adjective in Japanese.

(7) Kirei-(na) onna
   beautiful-PRT woman
   ‘Beautiful woman’

Similar to the claims for Japanese, Korean too admits only a relative clause structure in the attributive position (Kim 2002). Traditionally, however, Korean has been analyzed as having adjectives but as seen in (8) the adjective is realized as a participial form and the relative clause marker attaches to the entire constituent.

(8) Ce [ e₁ yeppu-ess]-ten₁ yeça
   that [ pretty-PRT]-REL woman
   ‘that woman who used to be/was pretty’

Slave and Ika (belonging to the Athapaskan language family) also admit adjectives in the attributive position only if there is an obligatory copula present on the adjective.

2 Hajime Hoji p.c.
We have seen that languages can resist attributive modification for adjectives. In Malayalam and other Dravidian languages, neither can an adjective occur in the attributive position nor in the predicative position, to which I turn to presently.

3. The Basic Paradigm in Malayalam

There are two classes of roots in Malayalam that look adjectival, in that they participate in positions in which English would have an adjective - Class1 or Relativizing roots and Class2 or Nominalizing roots. A brief look into the history of these roots suggests that Class1 roots had a verbal origin (See Jayaseelan 2007) and could be deverbal (as suggested in Anandan 1985). Class2 roots are borrowed roots, mostly from Sanskrit.

(10) **Class 1** (-a ending relativized roots)


(11) **Class 2** (-am ending nominalized roots)


The forms belonging to Class1 roots all end in –a which is also the Proto-Dravidian relative clause marker derived from a shortening of the distal determiner aa ‘that’.

(12) a. pazhay-a ‘that which is old’ Tamil, Malayalam
    b. p-a ‘that which is old’ Kodagu, Todi
    c. par-a ‘that which is old’ Tulu
    d. hos-a ‘that which is new’ Kannada, Tulu
    e. pedd-a ‘that which is great’ Telugu

The idea that words found in the Class1 category are reduced relative clauses was first suggested by Anandan (1985). The Class2 roots are borrowed mostly from Sanskrit. Malayalam has a phonological restriction on the coda position of a syllable. The only sounds that can appear in this position are vowels and the bilabial nasal /m/ and the alveolar nasal /n/. Notice that the Sanskrit roots in Class2 mostly end in an obstruent. This phonological coda restriction entails that the nominal morpheme –am is employed to turn the Class2 roots into something more native-like. Malayalam also has a nominalizer atə which is used to nominalize only clauses or verbal elements as we will see below.

The Class1 roots can undergo nominalization with the nominalizing morpheme atə whereas Class2 roots cannot, since they are already nominals; neither can they be relativized using the -a marker.

---

3 Among the Dravidian languages, Malayalam borrowed the most from Sanskrit and Tamil resisted Sanskritization.

4 Even though Malayalam has the maximum number of nasals in any Indian language, only these two nasals can occur in the coda position.
In the next section, I will look at the distribution of the Class1 roots and Class2 roots looking specifically at the attributive and the predicative positions.

3.1. Distribution of Class1 and Class2 Roots

The Class1 and the Class2 roots display different syntactic behavior. Relativized Class1 roots can appear in attributive positions whereas Class2 roots can appear in attributive positions only with the mediation of a non-finite copula (uLL, the verb ‘to exist’) and the relative marker –ə.

(14) a.  valiyə kutti  [Class1]
    big    child
    ‘Big child’ (Lit: child being big)

  b.  santoshəam uLLa kutti  [Class2]
    happy    COP-REL    child
    ‘Happy child’(Lit: child (to whom) there being happiness)

This strategy is also found in other languages such as Wolof as reported by McLaughlin (2004) and Slave and Ika as discussed above. The relative clause marker in Wolof is /Cu/ where C is a noun class marker which shows concord with the noun.

(15) a.  xale bu rafet  Wolof
        child    REL    pretty
        ‘A pretty child’

  b.  xale bu xam
        child    REL    know
        ‘A child who knows’

The use of the non-finite copula ties in with the fact that relative clauses in Malayalam are non-finite (See Jayaseelan 2011 for a detailed analysis). A question begs itself at this point, why can’t the relative clause marker attach directly to the borrowed roots, i.e. why doesn’t the language allow words as in (16).

(16) a.   *pokk-a  ‘tall’

  b.  *santosh-a  ‘happiness’

If both Class1 and Class2 roots are identical then what makes the relative clause marker attach only to certain roots? Class2 roots as we saw before are borrowed roots. Moreover there is no prohibition in a word ending in a vowel (as Class1 illustrates). The answer lies in

Interestingly, this form is attested only in one usage- the ‘Happy birthday’ equivalent in Malayalam which is:

(i) santoshajanmadinam kutti-kkə
    happy born day child-DAT
    ‘Happy birthday to the child’

Presumably, this is because ‘santoshajanmadinam’ is a compound and the /m/ in the coda of ‘santosham’ is deleted.
the patterns exhibited by the two classes of roots. The morphology module admits both the relativization and nominalization as routes to realizing an adjectival meaning. The Class1 roots are deverbal, the Class2 roots are not. –a is always looking for a verbal element. Class2 roots are borrowed and upon borrowing has to undergo the nominal morphology prior to the –a suffixation.

The fact that –a is always looking for a verbal element is exemplified in the attributive position of the Class2 adjective, they always need the non-finite copula as support for the –a attachment. –a attaches to verbs and never to nouns.

(17) a.  [njaan ___ kaNT-a] kutti
     I see-REL child
     ‘The child that I saw’

b.  [[njaan ___ kaNT-u enna] ningal parayunn-a] kutti
     I see-PAST COMP you say-REL child
     ‘The child that you say that I saw’

In English, the relativizers move from an argument position to a non-argument position. The Malayalam –a is different in that respect. The behavior of –a is different from the English ‘who’ in that –a does not open up any argument positions. –a does not contribute anything semantically, since the root has already been changed into an <e,t> with a null verbal head in the case of Class1 and by the overt non-finite copula in Class2 as we will see in the next section.

Neither Class1 nor Class2 roots can appear as predicative modifiers unless they are nominalized. Class1 roots appear in relative clauses, which now modify a pronominal. This is seen by the number and gender marking on the relativized Class1 root that combines with the so-called equative copula, whereas Class2 roots can appear as the complement of an existential copula triggering the appearance of dative case on the subject/experiencer.

(18) a.  avan nalla-vaN aaNə
     he good-M.SG EQ
     ‘He is good’ (Lit: he is one being good)

b.  avaL nalla-vaL aaNə
     she good-F.SG EQ
     ‘She is good’ (Lit: she is one being good)

(19) a.  avan-ə pokkam uNTə
     he-DAT tall EX
     ‘He is tall’ (Lit: to him there is tallness)

b.  kutti-kkə dukham uNTə
     child-DAT sad EX
     ‘The child is sad’ (Lit: to the child there is sadness)

Class1 and Class2 appear in predicative positions with different copulas. While Class1 uses the equative copula aaNə, Class2 uses the existential copula uNTə. While in Class1 roots, there is the appearance of the nominalization morphemes adhering to the number and gender features of the subject and the subject appears in nominative case. Class2 are nominals to begin with, hence, there is no reappearance of the nominal markers found in (18). The strategy that is at use here is what I call the ‘possessive strategy’. The dative case in (19) exemplifies this overtly. The meaning is akin to saying the subject ‘he’ possesses ‘tallness’ or ‘height’. The ramifications of this proposal will be elaborated in the analysis section below. A summary of the facts is below:
(20) Table 1.0

<table>
<thead>
<tr>
<th></th>
<th>CLASS1: NATIVE ROOTS</th>
<th>CLASS2: BORROWED ROOTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Nominalization</td>
<td>$\sqrt{1} + at_\text{noml}$</td>
<td>$\sqrt{2} + am$</td>
</tr>
<tr>
<td>b. Attributive</td>
<td>$\sqrt{1} + a$</td>
<td>$\sqrt{2} + am$ $\text{EX}_{\text{non-finite}} + a$</td>
</tr>
<tr>
<td>c. Predicate</td>
<td>$\sqrt{1} + a + \text{AGR EQ}$</td>
<td>$\sqrt{2} + am$ $\text{EX;}$ $\sqrt{2} + am$ $\text{EX}_{\text{non-finite}} + a + \text{AGR EQ}$</td>
</tr>
</tbody>
</table>

4. The Analysis

My core proposal is that Dravidian never lexicalizes an adjective, in other words, an $A$ does not exist in the lexicon of Dravidian nor does it derive one in the syntax. The only primitive categories that do exist are Ns and Vs. As and Ps are always derived in the syntax-morphology interface. I assume the lexicon to contain only roots (similar to Halle and Marantz 1993, Marantz 2004, Borer 2003) as in the Distributed Morphology tradition. These roots are prototypical ‘property concepts’ and refer to kinds ($e^K$). An adjectival meaning is expressed by either a reduced relative clause structure (in the attributive position) or as a nominalization (in the predicative position). The two routes to the adjectival meaning are mediated by a possessive semantics.

Keeping the lexicon devoid of any inflectional morphology, i.e. the morphological particles exist in the functional lexicon but they are not attached to the roots, allows many one-to-many mismatches to surface only in the morpho-syntactic module. Derivations are syntactic and can be seen in additional functional structure which contribute to interpretation. The Class1 and Class2 roots start out as category-neutral expressions of type $e^K$. In the morpho-syntactic module they undergo complex derivational processes that enable them to function as words, thus word formation is always in the syntax.

4.1 Class1 Roots

Recall Class1 roots have been traditionally assumed to be deverbal (See Old Malayalam data in Jayaseelan 2007). They can only be merged in the complement position of a vP which has a null verbalizer head. The root first composes with the null head. I will argue that this is essential and the only way for the relative clause marker $-a$ to attach to the root. It can never combine directly with the root without the mediation of this extra functional layer. The $-a$ marker can only attach to verbal predicates. This also explains why $-a$ can never attach to Class2 roots directly, because they are nominals to begin with.

In the attributive position, the Class1 root can appear as a reduced relative clause. $-a$ is itself not an $A'$ operator but a morpheme on the verb that marks what argument has been relativized (See Caponigro and Polinsky 2008, Caponigro and Polinsky 2011). Crucially, relative clause markers such as ‘who’ in English moves from an already created argument position and in some languages there is a requirement that only the subject position can be relativized. However, the semantics of the relative clause marker in the Class1 roots cases is simply to make the Class1 verbalized root into a reduced relative clause. I assume the following semantics for the null verbalized element. This is a modification of the semantics of the possessive ‘ka’ in Ulwa (see Francez and Koontz-Garboden 2010).

(21) $\| \emptyset \| = \lambda x. \lambda \Pi. \Pi x. [\pi (x, \Pi)]$

(21) is the semantics of the null verbalizer for Class1 roots. $\Pi$ ranges over entities that have a kind reference, a.k.a roots. $\pi$ expresses the possessor relation. The verbalizing head
itself plays the role of one of those operators, essentially turning a kind into a property. The calculation proceeds thus:

(22) **Step 1:** Combination with the null verbalizer
\[ \sqrt{\text{nall}_{e}^{K} + \emptyset} \}_{v} = \text{nall}_{<e,t>} \]

**Step 2:** Combine with the relative clause marker
\[ \text{nall}_{<e,t>} + \text{-a}_{\text{rel}} = \text{nalla}_{<e,t>} \]

Note that the relative marker does not change the semantic type of the predicate, but allows for syntactic function as an attributive modifier. A verb cannot function by itself. The derivation is represented as tree diagrams below:

(23) **Step 1:** Combination with the null verbalizer

```
    vP  
   /   \ 
  V   \sqrt{good}  
      /  \  \  
    v   \emptyset
```

**Step 2:** Combine with the relative clause marker

```
    DP  
   /   \  
  xAP  NP  
      / \  /  
     FP  N'  N'  
        /  /  /  
       a  vP  v  \sqrt{child}  
          /  /  
         v  \emptyset
```

The fact that attributive Class1 roots are always reduced relative clauses is given further support by the absence of non-intersective readings. In English, (24) below is ambiguous between an intersective reading and a non-intersective reading (Siegal 1980) whereas the Malayalam counterparts in (25) only show an intersective reading suggesting they are actually reduced relative clauses.

(24) Olga is a beautiful dancer
    Reading 1: Olga is a dancer and Olga is beautiful
               **intersective**
    Reading 2: Olga is someone who is a dancer and her
dancing style is beautiful
               **non-intersective**

(25) Sita oru pazhaya nrithakkari aaNo
    Sita a old dancer EQ
    ‘Sita is an old dancer’ (= she used to dance and no longer does)
    Reading 1: Sita is someone who is an old dancer
               **intersective**
Adjectives inside English relative clauses behave like the Malayalam examples in (25) in displaying only the intersective reading. This lends support to the reduced relative clause analysis. Semantically, the null verbalizer converts the root into a predicate of type $<e,t>$. The semantic role of the $\neg a$ is only to make the predicate into a reduced relative clause. The syntax of the reduced relative clause is sketched below:

(26) **Reduced relative clause**

the [boy, $\neg a$ boy, being good]

\[ \text{LF: the } [[\lambda x \ [\text{boy, } x]]] \ [\text{being good}]] \]

\[ (\text{combine the two predicates by Predicate Modification}) \]

\[ t(\lambda x[\text{boy } (x) \ ^\pi (x, \text{good})]) \]

The predicative position, I noted, also requires a nominal. The verbalized roots cannot appear in this position without the help of additional nominal morphology – and as relative clauses they cannot appear there either. This nominalization is sensitive to the number and gender of the subject (cf. (18) and (19). Baker (2003) in analyzing predicative adjectives assume they check selectional features of the PRED head. Similarly, the predicative head [+PRED] in Dravidian is marked for nominal features and these features have to be checked off by the operation [AGREE]. The appearance of the nominal features is only a reflex of the checking operations.

Commonsensically, it is plausible for the $\neg a$ marked root to appear in the predicative position since it is already a predicate however, syntactically relative clauses are not standalone predicates. Moreover, the clause structure of Dravidian is very restricted and conservative (see Jayaseelan 2011 for a recent discussion of this idea). Jayaseelan (2011) takes this conservativity to be seen as the inability to “hive-out” positions in the clause architecture. Supposing that what I have said is on the right track and there are indeed no adjectives in Dravidian, then we expect only a nominal element as the complement of the verb. The inability of the $\neg a$ marked root to appear as the complement of the verb suggests that only a nominal can appear in the complement position of the verb. If this theory is correct, it makes two predictions which we find borne out in Dravidian:

a. Comparative constructions formed with Class1 roots are always nominalized, since there can only be nominal comparatives. Presumably verbal comparatives should also be allowed. Adjectival comparatives should be missing.

b. Secondary predications of the kind found in English should not be possible.

I have established why the Class1 roots have to appear with nominal morphology in the predicative position. I will now proceed to my assumptions on AGREE and the presence of the nominalization morpheme. The model I assume is closest to recent modifications of Chomsky’s original AGREE model proposed in Frampton and Gutmann (2006) and Pesetsky and Torrego (2007), henceforth FG and PT. They propose a feature-sharing model where the probe can evaluate and check features of the probe by multiple AGREE. In looking at Icelandic data which show agreement on the pronoun as well as the matrix participle, the FG’s analysis would entail the feature sharing mechanism whereby the participle first agrees with the pronoun and then subsequently the matrix $\nu$ can check and assign case to the pronoun and this case is shared with the participle. The pronoun is in some sense linked to the participle.

For Dravidian, the Class1 roots are realized as participials in that they are reduced relative clauses having very little structure. I already noted the requirement of the Dravidian
verbal predicate to have a nominal in its complement position. Adopting FG and PT’s feature sharing approach enables us to explain why the ClassI root in the predicative position always appears with a nominal marker sensitive to gender and number. I will assume a [+PRED] head that has {N, G} features. The probe on the [+PRED] head is looking for some element to saturate its feature. The reduced relative clause cannot saturate this without the help of the nominalization morpheme- *van, *val, and *to. These rudimentary agreement markers have phi-features that can saturate the [+PRED] head.

(27) a. \([\text{PRED} \text{nalla} [\text{VP} \text{ COP}]]\)

b. 

[Diagram]

The [PRED] head’s features have to be satisfied by the element in the complement position. This is done by the feature sharing mechanism since this is reflected in the case assignment patterns as well. The nominative case is a feature assigned as a result of the feature sharing by the probe and goal. Thus, the subject in the case of the ClassI predicative constructions is assigned nominative case (which is always null marked). The feature sharing mechanism sketched here is different from Baker’s agreement checking story for the lack of adjectives in the attributive position in Japanese. Baker’s (2003) story relies on feature checking where the adjective has to be specified for uninterpretable features in order for the feature checking to happen. In this case, it is not about checking off uninterpretable features, rather the need to share features in connection with the fact that only a nominal can appear in the complement position of the predicative head that triggers the feature checking mechanism. In the absence of a lexical category of adjectives, the fact that the complement of the [PRED] head is a nominal is indeed not surprising. The derivation is below:

(28) a. \([\text{CP} [\text{TP} \text{ John-ø [PRED } \text{ nalla-van [COP aaNa]]]}]]\)

b. 

[Diagram]

I am assuming the Spec of the VP to host the xAP projection as shown in (23: Step2). The ClassI root first combines with the null verbalizer which then combines with the relative clause head. This is then merged into the Spec of the VP. The probe on the [PRED] head sends out the probe which then agrees with the xAP. The movement of the root from lower down in this projection to the Spec of the xAP (as seen in 23:Step2) is motivated precisely because of this AGREE relation. Only the edge of the projection is visible to the probe.
Depending on the PNG features, the inflection on the relativized root changes. This thus creates a nominal predicate.

4.2. Class2 Roots

Class2 roots are borrowed roots. These roots normally end in a stop or a fricative. Malayalam coda is restricted in that only a bilabial nasal, an alveolar nasal or a vowel can appear in the coda position. The roots are nominalized after being borrowed with the nominalizing morpheme –am. Note that the other nominalizing morpheme ‘atə’ is not available for Class2 roots, because ‘atə’ is a true nominalizer and can appear only on verbal and clausal elements whereas –am is a nominal morpheme and appears in other non-borrowed nominal roots as well.

(29) a. pazhəm ‘banana’  b. veLLam ‘water’

A question that comes up automatically is why the Class2 roots cannot undergo relativization, unlike the Class1 roots. The relative clause is not attached directly to the root, it attaches to the null verb. Therefore the null v cannot attach to Class 2 roots. Class2 roots, unlike Class1 roots, thus start out as nominals. There is no syntactic restriction prohibiting the attachment of the relative clause marker onto the borrowed root, the constraint is purely phonological. Borrowed roots have to undergo the nominalization because of the coda restriction. Recall that the relative clause marker, I argued in § 4.1 can attach only to a verbal predicate. Thus, the only way for the relative clause marker to attach to the Class2 roots is by converting the Class2 roots into a verbal predicate. However, this is not the strategy that is commonly used for borrowing.

The Class2 roots, similar to the Class1 roots start out as eK. Addition of the nominal morpheme –am turns the roots into e. The nominal morpheme is the head of an nP. The semantics of the word created by the –am suffixation behaves like any other nominal in the language. Given that Class2 roots end up as nominals and not as predicates, attributive modification is not possible at all. The only way to use a nominal predicate in an attributive position is to employ a copula. Malayalam uses the non-finite existential copula uLL for this purpose. The non-finite copula turns the nominal predicate into an <e,t> to which the relative clause marker –a can then attach to. The use of the non-finite copula is consistent with the view that relative clauses are non-finite in Dravidian. This non-finite copula is the overt form of the null verbalizer employed for the Class1 roots. Thus, Class2 roots have the morphology spelt out overtly whereas Class1 roots only have it covertly. The calculation proceeds thus:

(30) a. **Step 1**: Combination with the nominal morpheme

\[ \sqrt{pokk_e^K + -am}_n = pokkam_e \]

b. **Step 2**: Combine with the non-finite copula

\[ [pokkam_e + uLL]_v = [pokkam uLL]_{<e,t>} \]

c. **Step 3**: Final merge with the relative clause marker

\[ [pokkam uLL_{<e,t>} + -a]_{rel} = pokkam uLLa_{<e,t>} \]

Note that –a is merging with a non-finite verbal head in this case the non-finite copula ‘uLL’. The syntactic derivations are show below:
In the predicative position, unlike the Class1 roots, there are two strategies that Class2 roots employ. Either the derivation can proceed similar to the Class1 roots, whereby the Subject gets nominative case and the Class2 predicate gets nominalized or the Class2 root can stay as in Step 1 of (31) and the subject can receive dative case. I will appeal to the feature sharing mechanism and show that this much-discussed “dative experiencer” construction receives a simple explanation under this theory. It is a by product of feature sharing as well as the semantics of the possession which I have discussed under the rubric of what it means to be in an attributive or predicative position.

When the Class2 root remains as a relative clause, it has to appear with the nominalizing morphemes pertaining to person, number, and gender. This is similar to the predicative position in a Class1 root. The [PRED] head is marked for features and there has to be obligatory feature sharing between the head and the Class2 root. The entire relative clause is nominalized. This would seem a bizarre strategy considering that the Class2 root was a nominal after the first morphological merge unlike the Class1 root (which ends up verbal). However, there is an option to use the nominal first created after the morphological merge as well which is the experiencer dative construction.

In the case of the first strategy, the [PRED] head sends down the probe which then agrees with the relative clause upon the affixation of the nominal morpheme. The AGREE relation shares the features on the NP and the relative clause which gets the nominalized forms.

(32) a. [CP] [TP] John-ø [PRED [PRED’ pokkam uLLa-van [COP aaNә]]]
In the second strategy, the VP combines with the nominal belonging to Class2. The PRED head then sends the probe to check the features on the nominal. Since it is already a nominal there is no further nominalization required and hence no addition of the nominalizing morphemes. This feature sharing results in the case assignment as well. Nominative case is assigned to the predicate. And thus, the “subject” cannot be assigned nominative case and the only other option is to assign dative.

5. Conclusion

In this paper, I have shown that Dravidian does not have an adjective category either lexically or derivationally in the syntax. The two routes taken to derive adjectival-like meaning are relativization and nominalization. I assume the lexicon comprises of roots which combine with different v heads and n heads in the syntax to derive the relevant structures. These roots start out as kinds \(e^k\). I will end with a section on a special strategy used by Kannada in creating adjectival-like structures and another section on adjectives in English where they cannot appear prenominally.

5.1. Kannada Adjectives

Jayaseelan (2007) notes that in Kannada the normal way of expressing Class2 roots is by using the dative case on the root.

(33) raama udda-kke idd-aane
    \(\text{raama tall-DAT be-3SG.MASC}\)
    ‘Rama is tall’

The nominative counterpart of (33) has a specialized usage and is used only to indicate a special context such as ‘Rama is tall enough to join the army’.

(34) raman-ige udda ide
    \(\text{raman-DAT height is}\)
    ‘Rama is tall’

This pattern is interesting and I will leave it for future research to explore if the other Dravidian languages use this pattern to arrive at similar meanings.
5.2. English Adjectives that cannot Appear Prenominally

In English, there is a subclass of adjectives referred to in the literature as a-adjectives beginning with the syllabic shwa that resist prenominal attributive modification.

(35) ??the asleep boy

(36) the sleepy/absurd/active/tall boy

Historically it is known that adjectives such as ‘asleep’ were prepositional phrases in Old English (Simpson and Weiner 1989). They are derived from a noun incorporating into a preposition.

(37) a. John is on/at sleep Old English
    b. John is asleep Modern English

Now suppose this suggests that adjectives which have more structure cannot appear in attributive positions. This then predicts that Dravidian roots cannot attributively modify a noun since they always compose with v and n heads to form relativized or nominalized structures.

References

Caponigro, Ivano and Maria Polinsky. (2008) “Almost everything is relative in the Caucasus”. Semantics and Linguistic Theory (SALT) 18, Amherst, MA.


ELLIPSIS AND RECONSTRUCTION IN TOUGH INFINITIVES*

Troy G. Messick
University of Connecticut

1. Introduction

The derivation of the English tough-construction (TC), shown in (1), has been an issue of contention since the advent of generative grammar.

(1) a. The book was easy to read.

   b. The book was easy [CP C [IP PRO to read t]]

   c. The book was easy [CP C [IP PRO to read the book]].

   d. The book was easy [CP Op C [IP PRO to read t]]

   e. The book was easy [CP t C [IP PRO to read t]]

The central issue in the debate over TCs is the relationship between the gap in the embedded infinitival clause and the subject of the matrix clause. There have been several proposals concerning the nature of this relationship ranging from: single step A-movement (Rosenbaum 1967, Postal 1971) (1b); an object deletion process (Akmajian 1972, Lasnik & Fiengo 1974) (1c); null operator movement (Chomsky 1977) (1d); and two step movement (Brody 1993, Hornstein 2001, Hicks 2009, Obata & Epstein 2012, Hartman 2012).

This paper makes two claims that add to the discussion of TCs: (1) TCs are governed by MaxElide. Ellipsis of the embedded VP of TCs patterns with ellipsis in VPs that contain an A’-trace and is governed by the constraint MaxElide (Merchant 2008). This novel data patterns with established tests for A’-movement: parasitic gaps (Chomsky 1982) and extraction out of double object constructions (Edmonds & Whitney 2006). This observation suggests that a derivation of TCs that solely involves A-movement (Postal 1971), a deletion process (Lasnik & Fiengo 1974) or does not leave a trace (Hartman 2012) cannot be maintained; (2) TCs exhibit (anti)-reconstruction effects similar to relative clauses. TCs display reconstruction effects with regards to Condition A of the binding theory, variable binding, and idiom chunks. TCs display anti-reconstruction effects with regards to Condition C of the binding theory and quantifier scope. These facts lead us to conclude that the TCs involve two-step movement; however, we depart from the analyses cited above in that I argue the second step of movement is not A-movement, but instead a movement similar to the one

* I would like to thank my colleagues at the University of Connecticut for creating a intellectually stimulating place to work. I would also like to thank the audience at GLOW in Asia IX at Mie University. I would like to especially thank Jonathan Bobaljik, Željko Bošković, and Susi Wurmbrand for discussion of the topics presented here. Any errors are my responsibility.
found in head-raising relative clauses (cf. Sportiche 2006). This argument is strengthened by the fact the two constructions are both sensitive to extraposition.

The paper is structured as follows: the second section investigates the first claim that TCs are governed by MaxElide and also reviews arguments un favor of TCs involving $A'$-movement; the third section investigates the second claim that TCs and relative clauses display the same (anti)-reconstruction effects; the fourth section sketches out a derivation of TCs that builds off the intuition that the mechanism behind TCs is the same as the one that underlies the derivation of relative clauses. The fifth section is the conclusion.

2. *Tough* Constructions Involve $A'$-Movement

2.1. Verb Phrase Ellipsis

While the VP complement of control (2a,b), raising (2c,d), and ECM (2e,f) infinitives can undergo ellipsis, eliding the VP complement of a Tough Infinitive results in a degraded utterance (3) (elided material is shown in <...>).

(2) a. John wants to win, but Mary doesn’t want to <win>.
   b. Kim isn’t sure that she can solve the problem, but she will try to <solve the problem>.
   c. John doesn’t like math, but Mary seems to <like math>.
   d. The printer works but the copier doesn’t seem to <work>.
   e. They say Mary doesn’t like raisins, but Bill believes her to <like raisins>.
   f. John wants (for) his team to win and Jill wants (for) her team to <win>.

(Wurmbrand 2011:3)

(3) a. *?With a cheat sheet, History tests are possible to pass, but even with one, chemistry tests are not possible to <pass t>.
   b. *? On most diets, fats are important to avoid, but on this one, carbs are important to <avoid t>.
   c. *On most days John is easy to please, but today, he is not easy to <please t>.
   d. *I know Federer is difficult to beat, but I want to know if Nadal is difficult to <beat t>, as well.

Interestingly, the expletive counterpart of TCs also license ellipsis of the VP complement of the infinitive, shown in (4).

(4) a. *?With a cheat sheet, it is possible to pass this test, but without one, it is not possible to <pass this test>.
   b. On most diets, it is important to avoid fats, but on this one, it is not important to <avoid fats>.
   c. On most days, it is easy to please John, but today, it is not easy to <please John>.

The question arises: why should this asymmetry exist? The ellipsis paradigm illustrated in (3) patterns with ellipsis of a VP that contains an $A'$-trace (Lasnik 2001, Fox & Lasnik 2003, Merchant 2008). $A'$-extraction out of a VP ellipsis site results in an ill-formed utterance.  

---

1 There is some variability in the judgments regarding raising and ECM infinitives. See Martin (2001).
2 There are some counterexamples to this claim when the $A'$-moved element c-commands an element in Focus (see Schuyler 2001).
Ellipsis and Reconstruction in *Tough* Infinitives (T. G. Messick)

(5)  
a. *They said they heard about a Balkan Language, but I don’t know which they did <hear about>.
b. *They attended a lecture about a Balkan language, but I don’t know which they did <attend a lecture about>.
c. ??They studied a Balkan language, but I don’t know which they did <study>.

(Merchant 2008:139)

These utterances are improved when a larger constituent is elided, as in sluicing, shown in (6).

(6)  
a. They said they heard about a Balkan Language, but I don’t know which <they heard about>.
b. They attended a lecture about a Balkan language, but I don’t know which <they attended a lecture about>.
c. They studied a Balkan language, but I don’t know which <they studied>.

(Merchant 2008:139)

To account for the contrast between (5) and (6), Merchant (2008) proposes *MaxElide*.

(7)  
\textit{MaxElide} (Merchant 2008: 141)

Let XP be an elided constituent containing an A’-trace. Let YP be a possible target for deletion. YP must not properly contain XP ($XP \not\subset YP$).

This constraint would explain the contrast between (5) and (6) because in both sets of examples there is an A’-trace in the VP. The target of sluicing (i.e. IP) would contain the target of verb phrase ellipsis, and thus deletion of the VP would violate *MaxElide*.

Turning to TCs, with *MaxElide*, we are able to account for why the utterances in (3) are ill formed, but when the matrix predicate is elided the utterances are improved.

(8)  
a. Without a cheat sheet, history tests are possible to pass, but even with one, chemistry tests are not <possible to pass t>.
b. On most diets, fats are important to avoid, but on this one, carbs are <important to avoid t>.
c. On most days, John is easy to please, but today, he is not <easy to please t>.
d. I know Federer is difficult to beat, but I want to know if Nadal is <difficult to beat t>.

(i) I know which books ABBY read and which ones BEN did.

(Merchant 2008:140)

3 Takahashi & Fox (2005) define MaxElide differently and it has the attractive quality of not being limited to A’-traces. However, their definition appears to possibly overgenerate examples, such as (i)

(i) ??Ben knows who she invited, but Charlie doesn’t know who <she invited>.

(Merchant 2001:58)

4 Lyn Tieu and Scott Anderbois (p.c.) raise the concern that sentences such as (i) violate a constraint like MaxElide yet seem well formed

(i) On most diets, fats are important to avoid, but on this one, carbs are important.

This example at first glance appears to be an instance of ellipsis, but it is important to note that the second constituent in the coordination is a well formed sentence on its own

(ii) Carbs are important

Suggesting that (i) is not a case of ellipsis. This becomes even clearer when you look at predicates that cannot “stand on their own”.

(iii)  
a. *Mary is easy.
b. *John is not easy to please but Mary is easy
Again, there is an A’-trace in the embedded VP and the matrix predicate contains the VP and is possible target for ellipsis, so deletion of the embedded VP would violate $MaxElide$ because it contains an A’-trace and it is properly contained within another possible ellipsis target.

It is important to note that $MaxElide$ as defined only targets A’-traces and does not target instances of A-movement. Merchant (2008) provides evidence for this with (9).

(9) a. Someone solved this problem.
   b. i. Who?
      ii. Who did?

(Merchant 2008:143)

Both VP ellipsis and sluicing are available in (9) because the trace in the VP is one of A-movement to the specifier of IP, so deletion of VP would not violate $MaxElide$. Also note that both the matrix and embedded predicates can be elided in control, raising and the expletive counterpart of TC infinitives shown in (10).

(10) a. John wants to win and…
   …Bill wants to <win> as well.
   …Bill does <want to win> as well.
   b. John is likely to run for president and…
   …Bill is likely to <run for president> as well.
   …Bill is <likely to run for president> as well.
   c. On most diets, it is important to avoid fat, but on this one…
   …it isn’t <important to avoid fat>.
   …it isn’t important to <avoid fat>.

No $MaxElide$ effect is found because again there is no A’-trace in the embedded VPs.

The analysis presented in this section crucially assumes that TCs involve A’-movement. In the two following subsections, I summarize two well-established arguments in favor of TCs involving A’-movement.

2.2. Parasitic Gaps

It is well established that constructions that contain A’-movement, such as $wh$-movement (11a) and relativazation (11b), license parasitic gaps.

(11) a. What did Bill read before filing $pg$?
   b. [The paper] that Bill read before filing $pg$.

As the examples in (12) show, TCs also share this ability. Importantly, a similar example that only involves A-movement (12c) does not have the ability to license parasitic gaps.

(12) a. [This book], is hard to buy without reading $pg$.
   (Chomsky 1982:56)
   b. [Lloyd Webber Musicals], are easy to condemn without even watching $pg$.
   c. *[Lloyd Webber Musicals], are likely to be condemned without anyone even watching $pg$.
   (Hicks 2009: 542)
2.3. Double Object Extraction

Finally, another well-established test for A’-movement is extraction out of double object constructions. It has been shown that A’-movement of the Recipient in double object constructions is disallowed. This is observed in *wh*-movement (13a), heavy NP-shift (13b), relativization (13c), topicalization (13d) and cleft constructions (13e).

(13) a. *Who did you give t₁ a present?⁵
    b. *John sent t₁ a letter [every musician in the orchestra].
    c. *These are the girls that the fool bought t₁ a gift.
    d. *?John, Mary said that she gave t₁ a gift.
    e. *It is Bill that John gave t₁ a book.

(Larson 1988:354)  
(Marantz 1993:133)  
(Larson 1988:355)  
(Citko 2011:138)

And again, we find that TCs pattern the same way, as shown in (14).

(14) a. *John is tough to give t₁ a present.
    b. *Kids are easy to tell t₁ a story.

(Edmonds & Whitney 2006:94)

Again, A-movement, such as passivization, (and A-movement followed by A’-movement) is allowed out of these constructions, as shown in (15).

(15) a. John was given t₁ a present.
    b. Who was given t₁ a present?

To summarize the main points of this section: ellipsis in the infinitives of TCs is governed by the constraint MaxElide. This is because TCs involve A’-movement (cf. control and raising infinitives). The claim that TCs involve A’-movement was further substantiated by known facts concerning parasitic gaps, and double object extraction. The novel data concerning MaxElide thus provides a new argument against proposals that posit TCs as a single instance of A-movement (1a) (Rosenbaum 1967, Postal 1971) or a deletion process (1b) (Akmajian 1972, Lasnik & Fiengo 1974). It can also be used as an argument against recent proposals that state that movement in TCs does not leave a trace (Hartman 2012).

In the next section, the reconstruction and anti-reconstruction paradigms for TCs will be reviewed. It will be shown that many of the effects found in TCs mirror those found in relative clauses, suggesting that the mechanism that underlies both constructions is the same.

3. (Anti-)Reconstruction in Tough Constructions

3.1. Reconstruction

TCs exhibit reconstruction effects in the domains of Condition A, variable binding, and idiom chunks. These effects dovetail with those exhibited by relative clauses. We will investigate each of these effects in turn.

---

⁵ In certain British English dialects, (12a) is improved when the Theme is definite.

(i) %Who did John give the book?
3.1.1. Condition A

TCs license anaphors in subject position, as shown in the examples in (16). If these anaphoric expressions fall under Condition A, then it would be necessary for the subject to reconstruct to a position that is c-command by the experiencer (i.e. the embedded clause).

(16) a. Getting herself\textsubscript{i} arrested on purpose is hard for me to imagine Betsy\textsubscript{i} being willing to consider.  
   (Postal & Ross 1971; as cited in Lasnik & Flengo 1974:540)  
   b. Pictures of himself\textsubscript{i} are difficult for every photographer\textsubscript{i} to ignore.  
   (Hicks 2009: 552)  
   c. This aspect of herself\textsubscript{i} is easy for Mary\textsubscript{i} to criticize.  
   (Pesetsky 2012:22)  
   d. Interest in each other\textsubscript{i} was difficult for [Mary and Liam]\textsubscript{i} to rekindle.

Relative clauses follow the same pattern, as shown in (17).

(17) a. The portrait of himself\textsubscript{i} that John\textsubscript{i} painted is extremely flattering.  
   b. The interest in each other\textsubscript{i} that John and Mary\textsubscript{i} showed was fleeting.  
   (Schachter 1973:32-33)

3.1.2. Variable Binding

Both TCs (18) and relative clauses (19) allow for reconstruction for variables to be bound by quantifiers.\textsuperscript{6} Again, to receive a bound interpretation the variable must reconstruct to be c-commanded by the experiencer.

(18) a. Pictures of his\textsubscript{i} friends are hard for every photographer\textsubscript{i} to sell.  
   (Sportiche 2006:8)  
   b. Friends of his\textsubscript{i} are easy for every photographer to shoot.  
   (Hicks 2009:554 n. 12)  
   c. The argument that his\textsubscript{i} student cheated is difficult for every professor\textsubscript{i} to dispute.  
   (Takahashi 2011:1)  
   d. The death of her\textsubscript{i} husband is difficult for every widow\textsubscript{i} to get past.

(19) a. The picture of his\textsubscript{i} mother that every soldier\textsubscript{i} kept wrapped in a sock was not much use to him.  
   b. John generally has an opinion of his\textsubscript{i} book that every novelist\textsubscript{i} respects.  
   (Bhatt 2002:52)

3.1.3. Idioms

Idiom chunks can appear as TC subjects, as shown in (20) and the NP head of relative clauses, as shown in (21). Under the assumption that an idiom must be together at LF to receive an idiomatic reading then these examples also provide evidence for reconstruction into the embedded clause.

(20) a. The hatchet was difficult to bury.

\textsuperscript{6} The inability of D heads, as in (i), to reconstruct will be discussed in section 3.2.2.  
(i) ??His\textsubscript{i} mother is easy for every boy\textsubscript{i} to talk to
b. Headway was easy to make.  
(Hicks 2009:554)
c. Pictures are easy to take with my new camera.

(21) a. The headway we made was satisfactory.
b. The careful track that she’s keeping of her expenses pleases me.  
(Schachter 1973: 31-32)
c. The pictures that Liam took came out nicely.

To summarize, TCs show reconstruction effects for Condition A, variable binding and idiom chunks. These facts suggest that the TC subject moved out of the embedded clause into the matrix subject position, but not all test point to this conclusion, as we will see in the next section.

3.2. Anti-Reconstruction Effects

TCs exhibit Anti-Reconstruction effects in the domain of Condition C and quantifier scope.

3.2.1. Condition C

Both subjects of TCs (25) and heads of relative clauses are able to circumvent Condition C effects. Condition C is assumed to hold throughout the derivation and so in order to circumvent the effect, there should not be a trace of the R-expression that is c-commanded by the experiencer.

(25) a. Pictures of John$_i$ are hard for him$_i$ to ignore.
b. *It is hard for him$_i$ to ignore pictures of John$_i$.
c. John$_i$’s mother is difficult for him$_i$ to please.
d. *It is difficult for him$_i$ to please John$_i$’s mother.

(26) a. I respect any depiction of John$_i$ he$_i$’ll object to.  
(Sauerland 2003:12)
b. The pictures of John$_i$ that he$_i$ likes are on his wall.

However, when an R-expression is placed in an idiom chunk, the idiomatic reading is lost in both TCs (27) and relative clauses (28), suggesting that reconstruction is not possible when it would violate Condition C.

(27) a. Pictures of himself$_i$ are easy for John$_i$ to take  (Idiom: √  Literal: √) 
b. Pictures of John$_i$ are easy for him$_i$ to take  (Idiom: *  Literal: √)

(28) a. the picture of himself$_i$ that Bill$_i$ took  (Idiom: √  Literal: √)
b. the picture of Bill$_i$ that he$_i$ took  (Idiom: *  Literal: √)  
(Sauerland 2003:19)

3.2.2. Quantifier Scope

In previous investigations of TCs, it has been found that subjects of TCs do not show reconstruction effects for quantifier scope. The TC does not appear to be able to take scope under the matrix predicate, as shown in (29).
(29) a. Nothing is hard for Melvin to lift.  
   \((\text{Nothing} > \text{hard}, \ast \text{hard} > \text{nothing})\)

   b. Few girls would be difficult for Jim to talk to.  
   \((\text{Few} > \text{difficult}, \ast \text{hard} > \text{few})\)  
   (Postal 1974: 356, 224)

   c. Many people are easy to talk to.  
   \((\text{Many} > \text{easy}, \ast \text{easy} > \text{many})\)  
   (Epstein 1989:651)

Even when combining an anaphor with the quantifier does not force reconstruction of the quantifier below the matrix predicate.

(30) a. No picture of himself is easy for Bill to ignore.  
   \((\text{No} > \text{easy}, \ast \text{easy} > \text{no})\)

b. Many of each other’s pictures are hard for [Mary and Bill] to ignore.  
   \((\text{many} > \text{hard}, \ast \text{hard} > \text{many})\)

The inability of quantifiers to reconstruct seems to be a piece of a larger paradigm: D heads do not reconstruct in TCs. This is illustrated in (33), where it is shown that possessive pronouns in D cannot reconstruct.

(33) a. *His friends are easy for every photographer to shoot.  
   (Hicks 2009: 552)

b. *Her work is hard to convince [every woman in the group] to share.  
   (Rezac 2006: 301)

This again patterns with relative clauses, as shown in (34).

(34) *His pictures that every boy loves are on the dresser.

To summarize the main points of this section: subjects of TCs show reconstruction effects with Condition A, (non-D head) variable binding, and idiom chunks. Suggesting that the subject moves out of the embedded clause. Subjects of TCs have the ability to circumvent Condition C, but not when placed within idiom chunks. The inability of quantifiers to reconstruction seems to be part of a larger paradigm that disallows D-heads to reconstruct. These characteristics are similar to those in relative clauses.

In the next section, I review some of the literature concerning relative clauses and show that TCs share even more in common with that construction. I then sketch out a proposal for TCs built upon proposals used for relative clauses.

4. **Tough Constructions & Relative Clauses**

As shown in the previous sections, TCs and relative clauses behave similarly in several domains, and thus suggest that the same mechanisms should be used to derive both constructions. Following Sauerland (2003), Hulsey & Sauerland (2006) and Hackl & Nissenbaum (2012), I assume that English relative clauses are in fact ambiguous between a raising structure (35a) and a matching structure (35b). In the raising structure there is only one NP head of the relative clause that undergoes A′-movement, but in the matching structure, there are two distinct heads: an internal head that undergoes A′-movement to the specifier of CP and external head that is the complement to D. The internal and external head must match (modulo vehicle change) and the internal head undergoes a mandatory ellipsis process.
Ellipsis and Reconstruction in *Tough* Infinitives (T. G. Messick)


(Sauerland 2003:4)

Relative clauses that require reconstruction (e.g. for Condition A, variable binding or idiom chunks) are unambiguously raising, while relative clauses that must not reconstruction (e.g. to circumvent a Condition C effect) are unambiguously the matching structure. All other relative clauses are ambiguous between the two structures.

The question now becomes are TCs similarly ambiguous. And the answer appears to be no. Hulsey & Sauerland (2006) show that the two relative clause structures behave differently with regards to extraposition. Raising structures do not allow extraposition (36), while matching structures do (37).

(36) a. I saw the picture of himself that John liked.
   b. *I saw the picture of himself yesterday that John liked.
   c. Mary discovered the book about himself that Bob wrote.
   d. *Mary discovered the book about himself yesterday that Bob wrote.
   e. *Mary praised the headway last year that John made.
   f. *I was shocked yesterday by the advantage that she took of her mother.

(37) a. I saw the picture of Clinton yesterday that John liked.
   b. Mary discovered the book about Rome yesterday that Bob wrote.

(Hulsey & Sauerland 2006:115)

In (36), the relative clauses are unambiguously raising since they require reconstruction for Condition A or because they contain an idiom chunk and the construction does not allow extraposition, but in (37), the constructions do not require raising and such do allow for extraposition. Heycock (1994) and more recently Bruening (2012) notice similar effects in TCs, but unlike relative clauses, all cases of TCs seem susceptible.

(38) a. *The papers were tough for me yesterday to grade.
   b. It will be tough tomorrow to get an audience with the Pope.
   c. *The Pope will be tough tomorrow to get an audience with.
   d. It was very hard last year to give up sugar.
   e. *Sugar was very hard last year to give up.
   f. It is enjoyable in the summer to eat strawberries.
   g. *Strawberries are enjoyable in the summer to eat.

(Bruening 2012:2-3)

(39) a. It was difficult on weekends to make headway.
   b. *Headway was difficult on weekends to make.
   c. It was hard last year to bury the hatchet.

d. *The hatchet was hard last year to bury.

It does not seem to matter whether the subject needs to reconstruct (39) or not (38), extraposition is not allowed in TCs, suggesting that TCs behave like raising relative clauses. The current two step proposals for TCs (Brody 1993, Hornstein 2001, Hicks 2009, Obata & Epstein 2012, Hartman 2012) do not seem fit to account for these facts, since the second step in those derivations is always A-movement, and as Bruening (2012) shows, extraposition does not interfere with A-movement, as shown in (40).

(40) Ruprecht seems in meetings to be a masterful commander.  
(Bruening 2012:3)

Based on these facts, I propose that second step in TCs is not A-movement as the previous work assumes, but is more similar to the second movement in the derivation of raising relative clauses. The departure is not as radical as it first sounds. For independent reasons, Hicks’s (2009) “smuggling” derivation of TCs looks strikingly similar to the raising analysis of relative clauses, as shown in (41) (cf. (35a)).

\[ \text{(41) The book was easy } [\text{CP } [\text{Op the book}] \text{ C } [\text{IP PRO to read } [\text{Op the book}]]] \]

The derivation in (41) would capture almost all the behaviors of TCs discussed; however, it still cannot account for the lack of reconstruction for scope. Hicks suggests:

“…an NP constituent of an A-moved DP may optionally reconstruct, whereas the D cannot. The D head is what determines scope relations, so the absence of reconstruction of D accounts for only the surface scope reading being available in TCs [Tough Constructions]”

(Hicks 2009:553).

An alternative to this account would be to take the parallelism between TCs and relative clauses all the way. In relative clauses, all that undergoes movement is the head NP. If the same were true in TCs, then the lack of reconstruction of the D head would be expected because there would be no trace of D in the embedded clause only the trace of NP (cf. Sportiche 2006). The derivation would then look like the one in (42).

\[ \text{(42) The book was easy } [\text{CP } [\text{Op book}] \text{ C } [\text{IP PRO to read } [\text{Op book}]]] \]

---

8 Hartman (2009) notices the same effect with experiencers.

(i) a. It is important to Mary to avoid cholesterol.
    b. *Cholesterol is important to Mary to avoid.

(Hartman 2009:390)

9 (39) is a simplified version of Hicks’s derivation, which actually has movement to the embedded SpecvP before movement to SpecCP.
5. Conclusion & Future Research

Before I conclude, I would like to discuss a few avenues for future research. First, the proposal made here would force the explanation of the lack of Condition C effects to be different in TCs and relative clauses since Sauerland (2003) makes use of the matching analysis to circumvent the effect. A possible explanation may lie in the mechanism used in circumventing Condition C in raising constructions.

(43) Pictures of John, seem to him, to be on the table.

To account for this Takahashi & Hulsey (2009) argue for what they call “wholesale late merger”. A similar mechanism can be used for TCs.

The second avenue for future research concerns case assignment. It appears that the TC subject needs to escape accusative case marking in the embedded clause. Even more interestingly, it appears that Case must be assigned in the embedded clause as illustrated by the unacceptability of passive and unaccusative verbs in the embedded clause (noted by Postal 1990).

(44) a. *He was easy to be arrested.
   b. *He is easy to arrive.

Yet again, this appears to pattern with relative clauses, as Bhatt (2006:12) points out the trace in infinitival relative clauses also must receive case. This is illustrated by the unacceptability of an infinitive relative clause that lacks the case-assigning of

(45) a. *[a person], to be fond of t_i.
   b. [a person], to be fond of t_i.

Hicks (2009) suggests that the null operator that contains the TC subject is assigned the accusative case in TCs and that would also work in these relative clause constructions under the assumption that raising relative clauses involve a null operator. The NP would then have its Case feature unvalued until it moved to the matrix clause where it is valued nominative.

To conclude, this paper provided novel evidence that TCs involve A’-movement and that movement leaves a trace in the embedded clause. The (anti)-reconstruction effects displayed by TCs lead us to conclude that there is a second step in TCs out of the embedded clause, but we depart from most mainstream two-step approaches in assuming that the second step in TCs is not A-movement, but is in fact more similar to the movement found in raising relative clauses. This assumption was further motivated by the fact that both TC and relative clause constructions are sensitive to extraposition, while A-movement is not. The lack of reconstruction of D heads in TCs can be accounted with a raising relative clause like derivation because all that would undergo movement would be the NP, and thus there would be no trace of D in the embedded clause for it to reconstruct to.

References

Bruening, B. (2012) “No such thing as “Defective Intervention”.” ms, University of Delaware.


1. Introduction

Two competing views of island constraints have emerged in recent work. In one, which we will call the accumulation view, islands result from the accumulation of several independently documented processing difficulties, such as the presence of a filler-gap dependency and an especially complex syntactic structure (cf. Kluender (1998, 2004), Hofmeister and Sag (2010)). In a sentence with an island violation, these difficulties are claimed to exceed the processor’s capacity and the sentence is perceived as unacceptable. In the second view, which we will call the disruption view, islands result from an otherwise unproblematic element that may disrupt a filler-gap dependency and render it illicit (cf. Ross (1967), Chomsky (1986), Rizzi (2004)). Bounding nodes and certain syntactic features, for instance, have been claimed to be of this type; they are not inherently bad, but they lead to unacceptability when they intervene in dependencies of certain types.

Though typically seen as competitors, these two approaches to island constraints are not in principle mutually exclusive. It could turn out to be the case that certain island phenomena are due to one and that others are due to the other, or even that a single island constraint stems both from the accumulation of difficulties that overtax the processor and from some element that is disruptive only when it intervenes in a dependency. It is also worth keeping in mind that although the accumulation view is typically associated with processing accounts and the disruption view with grammatical accounts, these associations are not logically necessary. It is not difficult to imagine small grammatical violations accumulating to result in very low acceptability or some element that disrupts the processing of a filler-gap dependency but does not produce any special processing burden on its own.

In this paper, we use the above two views of islands as a backdrop as we explore the role of finiteness in island constraints. Ross (1967) first noted that finiteness seems to make island domains even more resistant to extraction, particularly in the case of wh-islands, as in (1).

(1) This is a book which I can’t figure out
   a. ? [what to do about __ ]
   b. ?*[what I should do about __ ]

There is barely any perception of unacceptability in (1a), but there is much more in its finite counterpart in (1b). A similar phenomenon has been noted in subject islands, as in (2)

---

* We would like to thank the audience at GLOW in Asia IX for their many very useful comments on our oral presentation. The present work is part of an ongoing project on the role of finiteness in extraction. Earlier versions were presented at the Islands in Contemporary Linguistic Theory conference (Vitoria-Gasteiz) and the CUNY Sentence Processing conference (New York). We are grateful to the audiences there, as well as to Chris Barkley, Ivano Caponigro, Gabe Doyle, Simone Gieselman, Boyoung Kim, Robert Kluender, Roger Levy, Emily Morgan, and other members of the UCSD Experimental Syntax Lab. This work benefitted also from the invaluable help of research assistants Adrienne LeFevre and Michelle McCadden in the UCSD Experimental Syntax Lab.
Overall acceptability here drops considerably compared to (1), but the contrast between non-finite (a) and finite (b) is maintained. Finiteness in adjunct islands is less often discussed, but the same contrast appears to obtain (Szabolcsi (2006), Truswell (2011)).

The accumulation and disruption views of islandhood make very different predictions with regard to where we should see effects of finiteness on acceptability. Under the accumulation view, the fact that finiteness appears to lower acceptability in (1)-(3) would suggest that finiteness is one of the several factors, each intrinsically difficult for processing on its own, that contribute to the unacceptability of island violations (cf. Kluender (2004) and Hofmeister (2007)). We should then expect to see a finiteness effect very generally across syntactic contexts. Under the disruption view, on the other hand, there is no reason to think that finiteness is intrinsically difficult and we thus expect to see a finiteness effect when finiteness intervenes in some dependencies, but not more generally (cf. Cinque (1990), Manzini (1992) and Truswell (2011)).

The extent of the finiteness effect can therefore play a crucial role in determining whether island phenomena are best viewed in terms of accumulation or disruption. Being certain of the extent of the effect is problematic, however, since finiteness effects are notoriously subtle. We should anticipate, then, that whatever effects we may find will almost certainly be small, but we also want to be sure that they are real. Approaching this empirical challenge experimentally thus seems prudent, since this will allow us to detect small differences in acceptability with some confidence (cf. Cowart (1997), Myers (2009)).

The rest of this paper presents the results of six sentence acceptability experiments designed to address this question of the extent of the finiteness effect. The experiments themselves are discussed in section 2. Experiment 1 probes the existence of a finiteness effect in adjunct islands, Experiments 2-5 explore whether the effect is ultimately due to finiteness itself or simply to the presence of an overt argument that typically accompanies finiteness, and Experiment 6 looks for a finiteness effect in non-island complement clauses. General conclusions about these experiments and what they tell us about the accumulation and disruption views of islandhood are presented in section 3.

2. The Experiments

2.1. Experiment 1: Finiteness and Adjunct Islands

To begin, we must first make certain that our experimental method is sensitive enough to detect the often slight degradation in acceptability that has been claimed for finiteness in some environments. We choose adjunct islands for Experiment 1, since finiteness in this island type has not been extensively documented in the literature, so beyond anything else, our results here can serve a useful descriptive function.

Participants (N = 220) in this experiment judged the acceptability of sentences using a 7-
point scale, with 7 representing the highest acceptability and 1 the lowest. Experimental items were all questions containing an adjunct clause, using a 2x2 design that crossed question-type (yes/no vs. wh-) and finiteness of the adjunct clause (finite vs. non-finite). Sample stimuli are given in (4).

(4) a. Did the carpenter restore the table [after negotiating / he negotiated with the buyer]?
   b. Who did the carpenter restore the table [after negotiating / he negotiated with __ ]?

Participants saw 6 tokens of each type, counterbalanced following a Latin square design. In addition to the 24 experimental stimuli, each participant saw 40 filler items. Materials were presented in pseudo-randomized order.

The results from this experiment are displayed in (5).

(5) Results from Experiment 1

![Graph showing results from Experiment 1]

There was a highly significant main effect for question-type (p<0.001), with a mean rating of 5.70 for yes/no questions and 2.27 for wh-questions, reflecting the fact that the latter case involves extraction out of an adjunct island. Analysis of question type and finiteness reveals no significant effect for finiteness in yes/no questions (non-finite mean = 5.72, finite mean = 5.69, p=0.68), but a highly significant effect in wh-questions (non-finite mean = 2.38, finite mean = 2.16, p<0.001).

These results show that we were able to capture a finiteness effect in adjunct islands. Finiteness in the adjunct clause leads to a significant degradation in acceptability with extraction, as expected. Moreover, these results begin to cast doubt on an accumulation account of the finiteness effect, in that finiteness degrades only the extraction (wh-question) case, and not the yes/no question case.

---

1 The recent literature suggests that an n-point response method such as this is advisable for studies attempting to capture fine gradations in acceptability. See Bader and Häussler (2010), Fukuda, Goodall, Michel and Beecher (2012), Sprouse (2011), and Wescott and Fanselow (2008, 2011) for discussion.
2.2. Experiment 2: Overt Arguments and CNPC

The results of Experiment 1 are highly suggestive, but there is a confound that prevents us from attributing the finiteness effect in extraction cases to finiteness *per se*. That is, finiteness always co-occurs in our stimuli with the presence of an overt argument (e.g. *he* in the finite case in (4b)), as it does very generally in English, and this leaves open the possibility that what we are calling a finiteness effect is actually due to this overt argument. In fact, this possibility has some plausibility, since processing referents is known to pose a burden to the processor (Gibson (1998), Warren and Gibson (2002)) and this increased processing cost could reasonably be expected to depress acceptability.

In order to eliminate this confound, Experiments 2-4 hold finiteness constant while manipulating the presence/absence of the additional argument. We begin here with Experiment 2, which does this within the realm of the Complex Noun Phrase Constraint (CNPC).

Experimental items in this experiment were all questions containing a complex noun phrase, using a 2x2 design that crossed question-type (*yes/no* vs. *wh-*) and an additional overt argument in the embedded clause (presence vs. absence). The embedded clauses were all non-finite. Sample stimuli are given in (6).

(6) a. Does the principal like [the thought of *(the children)* learning subtraction]?
   b. What does the principal like [the thought of *(the children)* learning __ ]?

Participants (N = 209) saw 4 tokens of each type. In addition to the 16 experimental stimuli, each participant saw 94 filler items. Experiments 3 through 5, below, served as some of these fillers (the same participants participated in experiments 2 through 5). All other aspects of this experiment were as in Experiment 1.

The results from this experiment are displayed in (7).

(7) Results from Experiment 2

As expected, there was a highly significant main effect for question-type (p<0.001), with a mean rating of 5.61 for *yes/no* questions and 2.85 for *wh*-questions, reflecting the CNPC violation in the latter case. There was also a significant main effect for presence/absence of an additional overt argument (p<0.001). There was, however, no interaction between question-
Recall that within the complex noun phrases in these stimuli, finiteness is held constant (always non-finite), so the effect that we are seeing here must be due entirely to the presence of an overt argument. Note, however, that unlike the pattern seen in Experiment 1, the degrading effect here (i.e. that of the overt argument) is not specific to \textit{wh}-dependencies.

### 2.3. Experiment 3: Overt Arguments and Subject Islands

Experiment 3 is the same as Experiment 2 but with subject islands instead of CNPC. The 2x2 design crosses question-type (yes/no vs. \textit{wh}-) and the additional overt argument in the embedded clause (presence vs. absence). The embedded subject clauses were all non-finite. Sample stimuli are given in (8).

(8) a. Does the lawyer know that [(\textbf{the boy}) presenting the testimony] will convince the jury?  
   b. What does the lawyer know that [(\textbf{the boy}) presenting \_\_] will convince the jury?

All other aspects of the experiment are the same as in Experiment 2.

The results from this experiment are displayed in (9).

(9) Results from Experiment 3

![Graph](attachment:image.png)

As before, there was a highly significant main effect for question-type (p<0.001), with a mean rating of 5.38 for yes/no questions and 2.98 for \textit{wh}-questions, reflecting the subject island violation in the latter case. There was, however, no significant main effect for presence/absence of an additional overt argument (p=0.27), although numerically, the pattern is the same as in Experiment 2, with those conditions with overt arguments less acceptable (mean rating: 3.99) than those without (mean rating: 4.36). There was also no significant interaction between question-type and the additional overt argument (p=0.62).

Importantly, there is no effect here that is specific to the \textit{wh}-questions. In this way, the results above are like Experiment 2, where the additional overt argument had an overall effect, but unlike Experiment 1, where finiteness affected only the \textit{wh}-questions.
2.4. **Experiment 4: Overt Arguments and Complement Clauses**

Experiment 4 has the same design as Experiments 2 and 3 but with non-island complement clauses in place of complex noun phrases and subject islands. The 2x2 design crosses question-type (yes/no vs. wh-) and the additional overt argument in the complement clause (presence vs. absence). The complement clauses were all non-finite. Sample stimuli are given in (10).

(10) a. Does the architect want [(the contractor) to see the plans before Monday]?  
   b. What does the architect want [(the contractor) to see __ before Monday]?

All other aspects of the experiment are the same as in Experiments 2 and 3.

The results from this experiment are displayed in (11).

(11) **Results from Experiment 4**

![Graph showing results from Experiment 4]

There is no main effect for question-type here; the results for yes/no and wh-questions are virtually identical. This is as expected, given that complement clauses are not islands. There is, however, a highly significant main effect for the additional overt argument (p<0.001), with a mean rating of 5.54 for sentences with the overt argument and 6.14 for sentences without. There is no interaction between question-type and overt argument, which reflects the fact that the absence of an overt argument has a significant ameliorating effect on both question-types and does not seem to target wh-dependencies in particular.

We are now in a position to evaluate the results of Experiments 2-4 as a group. Recall that the purpose of these three experiments was to begin to disentangle the confound in Experiment 1, where we found significant degradation in sentences with wh-extraction out of adjunct clauses that had two properties: they were finite and they had an additional overt argument. Since it was not possible in that experiment to see which of these properties was driving the effect, Experiments 2-4 held finiteness constant and manipulated only the presence/absence of the overt argument. In Experiments 2 and 4, there was significant degradation when the overt argument was present, but this affected all question-types, not just wh-questions. In fact, in none of these three experiments did the overt argument effect distinguish between yes/no and wh-questions. It thus appears unlikely that this effect is responsible for the primary results in Experiment 1, where there is a significant difference.
between the two \(wh\)-question conditions, but not between the two \(yes/no\) question conditions.

### 2.5 Experiment 5: Finiteness and CNPC

Given the results of Experiments 2-4, we now predict that if we hold the additional overt argument constant and manipulate finiteness, we will get an effect that is specific to \(wh\)-dependencies. In Experiment 5, we thus use a 2x2 design that crosses question-type (\(yes/no\) vs. \(wh\)-) and finiteness (finite vs. non-finite) with complex noun phrases. The clauses within the complex noun phrases all contain an overt subject. Sample stimuli are given in (12).

12. a. Do many people believe [the idea that the squirrels buried / of the squirrels burying extra food under bushes]?
   b. What do many people believe the idea [that the squirrels buried / of the squirrels burying __] under bushes?

All other aspects of this experiment were as in Experiments 2-4.

The results from this experiment are displayed in (13).

13. Results from Experiment 5

As expected, there was a highly significant main effect for question-type (\(p<0.001\)), with a mean rating of 5.46 for \(yes/no\) questions and 2.61 for \(wh\)-questions. This makes sense, given the CNPC violation in the latter case, just as was seen earlier in Experiment 2. There is also an interaction between question-type and finiteness (\(p=0.02\)), and further analysis reveals a significant preference for non-finiteness only in the \(wh\)-questions, with a mean rating of 2.5 in the finite condition and 2.71 in the non-finite condition (\(p=0.003\)). In the \(yes/no\) questions, the preference goes the other way, with a mean rating of 5.62 in the finite condition and 5.29 in the non-finite condition (\(p<0.001\)).

For present purposes, the most important finding here is the preference for non-finiteness among the \(wh\)-questions. This result conforms to our expectations, given what we saw in Experiments 1-4, and it confirms the earlier conclusion that what appeared to be a finiteness effect in Experiment 1 truly was a finiteness effect. That is, Experiment 5 provides evidence, along with Experiments 2-4, that the confound in Experiment 1 can be disentangled and that it is finiteness, not the additional overt argument, which specifically affects \(wh\)-dependencies.
2.6 Experiment 6: Finiteness and Complement Clauses

Now that we have established that the finiteness effect is detectable experimentally (Experiment 1) and that it is due in fact to finiteness, and not to the overt argument that often accompanies finiteness (Experiments 2-5), let us return to the question raised in section 1 regarding the accumulation and disruption views of island constraints. Under the accumulation view, we expect finiteness to have a general degrading effect on acceptability, while under the disruption view, we expect it to have no particular effect normally, but to reduce acceptability specifically when it intervenes in a *wh*-dependency. The evidence that we have seen here so far very clearly supports the disruption view. As seen particularly in Experiments 1 and 5, finiteness degrades acceptability in *wh*-questions that violate island constraints, but not elsewhere. This result would be surprising if finiteness were simply one of many degrading factors that result in an island violation when they accumulate, since in that case, the finiteness effect should be detectable independently of the island environment.

To say that the disruption view of the finiteness effect appears to be on the right track leads immediately to the question of why finiteness would affect *wh*-dependencies in this way. One promising recent analysis is given by Truswell (2011), who proposes the Event Locality Condition shown in (14).

(14) The Event Locality Condition: Events form locality domains for *wh*-movement.

That is, *wh*-movement is felicitous only if the domain containing both the head and foot of the movement can be construed as a single event. This condition is relevant to our concerns here because finiteness often influences how events are structured within a sentence. Truswell notes, for instance, that finite adjunct clauses are normally interpreted as independent events, while non-finite adjunct clauses may allow for the possibility of being interpreted as part of the main clause event. Under this view, the finiteness effect in adjunct clauses that we observed in Experiment 1 derives from the idea that the entire sentence is more easily interpreted as a single event when the adjunct clause is non-finite. This event structure has no effect on acceptability in most cases, but it does when *wh*-movement is involved, given (14).

Truswell also notes that finiteness has no effect on the event structure of complement clauses (of bridge verbs). These clauses are interpreted as part of the main-clause event regardless of whether they are finite or non-finite. This then predicts that we will not observe a finiteness effect with *wh*-extraction out of complement clauses. We test this prediction in Experiment 6, where we use a 2x2 design crossing question-type (yes/no vs. *wh*) and finiteness (finite vs. non-finite) with complement clauses of bridge verbs. Participants (N=220) saw 3 tokens of each type, counterbalanced following a Latin square design. In addition to the 12 experimental stimuli, each participant saw 52 filler items. All of the sentences contain an additional overt argument as in Experiments 1 and 5. Sample stimuli are given in (15).

(15) a. Did the children believe [the guest *was* / *to be* bringing a cake]?
    b. What did the children believe [the guest *was* / *to be* bringing __ ]?

Participants used the same 7-point acceptability scale as in the other experiments.
The results from this experiment are displayed in (16).

(16) Results from Experiment 6

As is expected here, given the fact that this is a non-island environment, there was no main effect for question-type (p=0.11). There was, however, a significant main effect for finiteness (p=.01), with a mean rating of 4.96 for finite and 4.82 for non-finite. There was no significant interaction between question-type and finiteness (p=0.14).

Notice that although there is a main effect for finiteness in these results, it is the finite cases that are more acceptable than the non-finite ones. This is the opposite of what we have seen elsewhere, and moreover, the preference is not specific to \textit{wh}-questions. We thus do not find here the finiteness effect seen in island environments in Experiments 1 and 5.

This lack of a canonical finiteness effect is exactly what is predicted under Truswell’s (2011) analysis. Since complement clauses under bridge verbs always form a single event with the main clause, regardless of the finiteness of the complement clause, non-finite embedded clauses have no ameliorating effect on \textit{wh}-extraction. This contrasts with adjunct clauses (and presumably with clauses embedded in complex noun phrases), where finiteness encourages an interpretation in which the clause is an independent event. Under this interpretation, \textit{wh}-extraction violates the Event Locality Condition in (14).

3. Conclusion

We began this paper by contrasting the accumulation view of island effects, in which islands results from the co-occurrence of independently motivated processing difficulties, with the disruption view, in which islands result from an otherwise unproblematic element that disrupts the filler-gap dependency. We have seen that the finiteness effect in islands appears to be most consistent with the disruption view. Outside of \textit{wh}-extraction environments, finiteness does not have a uniformly degrading effect on acceptability, but within (island) \textit{wh}-extraction environments, it has a clear degrading effect. Moreover, this degrading effect occurs in exactly those environments predicted by Truswell’s Event Locality Condition. That is, the effect occurs in islands, but not in non-island complement clauses.

The accumulation view also receives support from some of the findings in this paper, however. Specifically, we have seen that adding an overt argument to a sentence consistently leads to a small decrease in acceptability and that this decrease obtains in \textit{wh}-questions as
well. The fact that this additional overt argument effect occurs in a variety of environments is most consistent with the accumulation view.

As discussed earlier, the accumulation view is typically associated with processing accounts, and the disruption view is typically associated with grammatical accounts. If we accept these associations, we then conclude that the additional overt argument effect is a processing effect and that the finiteness effect is a grammatical effect. This conclusion is especially tempting given that there are plausible analyses on each side: Warren and Gibson’s (2002) analysis of the processing of referents predicts the overt argument effect seen here, and Truswell’s (2011) analysis of event structure and wh-extraction predicts the finiteness effect in just the environments where we have observed it.

If this approach is on the right track, then island phenomena are at least partly a grammatical effect. This does not preclude the influence of processing factors as well, though it does suggest that islands are not reducible to such factors. In terms of the concrete phenomena explored here, the behavior of finiteness seems to require a grammatical account, but the addition of an overt argument may add a processing burden that further depresses acceptability. A full account of island phenomena, therefore, is likely to include both grammatical and processing factors that underlie the decrease in acceptability.

References

1. Introduction

This paper examines a Japanese elliptical construction which we call Japanese Why-Stripping (JWS). We will discuss properties of JWS in comparison with those of other constructions such as Japanese Sluicing. Sluicing is an elliptical construction that involves a fragmental wh-phrase as shown in (1a) (Ross 1969, Merchant 2001, a.o.). Japanese Sluicing (JS) (Inoue 1976), exemplified in (1b), also involves a wh-remnant and thus is traditionally considered to be the Japanese version of Sluicing, although as discussed below, we will analyse JS in a different way from English Sluicing.

(1) a. Mary bought something, but I don’t know what.
   ‘Mary bought something, but I don’t know what.’

JWS, on the other hand, is an elliptical construction with a wh-phrase naze ‘why’ followed by a non-wh remnant (e.g. ringo-o ‘apple-Acc’ in (2b); (2a) is an example of English Why-Stripping).1

(2) a. Mary bought an apple, but I don’t know why an apple.
   b. Mary-ga ringo-o kat-ta ga,
      Mary-Nom apple-Acc buy-Past but
      boku-wa [naze ringo-o ka] wakara-nai.
      I-Top why apple-Acc Q know-not
      ‘Mary bought an apple, but I don’t know why an apple.’

In this paper, we will focus on the similarities between JWS, JS, and the cleft construction (Cleft). Adopting Hiraiwa and Ishihara’s (2002) analysis of JS where JS and Cleft are both derived from the so-called ‘no da’ in-situ focus construction, we claim that JWS is also derived from this same underlying construction. We will also show that JWS is distinct from Multiple Sluicing such as (3) (Nishigauchi 1998, Merchant 1998, a.o.), where more than one wh-phrase survives ellipsis, despite the fact that both constructions involve

---

1 The term “Stripping” refers to an elliptical construction with a non-wh remnant (e.g. (i)).
(i) John ate an apple, but not an orange. (= but John didn’t eat an orange.)
Hoji (1990) treats examples like (iiB) as Japanese Stripping, which we will not discuss here.
(ii) A: John-ni-wa huransugo- ga totemo zyoozuni hanas-eru.
     John-Dat-Top French-Nom very well speak-can
     ‘John can speak French very well.’
B: Watasi-no musuko-ni mo da.
    I-Gen son-Dat also be-nonpast
    ‘My son, too.’
multiple remnants.


Instead, we will claim that the behavior of JWS is captured under the analysis where naze ‘why’ is base-generated in the left periphery (Ko 2005).

2. JS/JWS and Cleft

2.1. Two Lines of Analyses of JS

Before considering properties of JWS, let us review previous analyses of JS, which is a more well-studied construction in the literature. There are two major lines of analyses of JS. Takahashi (1993) claims that JS is derived via wh-movement and clausal deletion as shown in (4a). We will call his approach ‘the wh-movement analysis’ of JS. This analysis assumes that JS is basically the same construction as English Sluicing, which also involves wh-movement and clausal deletion according to authors such as Ross (1969) and Merchant (2001), among others (e.g. (4b)).

(4) a. Mary-ga nanika-o kat-ta ga, boku-wa Mary-Nom something-Acc buy-Past but boku-wa [ CP nani-o [ IP Mary-ga-ta-kat-ta-ka] wakara-nai. what-Acc Mary-Nom buy-Past Q know-not ‘Mary bought something, but I don’t know [CP what, [IP Mary bought]].

b. Mary bought something, but I don’t know [CP what, [IP Mary bought]].

Contra the wh-movement analysis, authors including Nishiyama et al. (1996) and Merchant (1998) claim that JS is derived from Cleft. (See also Hiraiwa and Ishihara’s (2002) related analysis, which will be discussed in Section 4). According to this ‘Cleft analysis’, the JS sentence in (1b) has the underlying Cleft structure in (5), whose focus is the wh-phrase, and whose presupposition clause undergoes deletion.

(5) Mary-ga nanika-o kat-ta ga, boku-wa Mary-Nom something-Acc buy-Past but boku-wa [CP Mary-ga-ta-kat-ta-no] ga nani-ko ka] wakara-nai. Mary-Nom buy-Past C-Nom what-Acc Q know-not ‘Mary bought something, but I don’t know what it is that Mary bought.’

The Cleft analysis has several advantages over the wh-movement analysis. The next subsection reviews the argument for the Cleft analysis.

2.2. The Arguments for the Cleft Analysis of JS

The following areas of Japanese syntax provide arguments for the Cleft analysis of JS: (i) the existence of the copula, (ii) the possibility of a pronominal subject, (iii) the optionality
of a Case-marker, (iv) a restriction concerning word order, and (v) island effects.\(^2\) The data below show a number of parallelisms between Cleft and JS in those areas, supporting the Cleft analysis of JS.

First, JS has an optional copula *da* “be” after the wh-phrase, as shown in (6).

(6) Mary-ga ringo-o kat-ta ga, Mary-Nom apple-Acc buy-Past but boku-wa [naze ringo-o da ka] wakara-nai. I-Top why apple-Acc be Q know-not

‘Mary bought an apple, but I don’t know why an apple.’

As (7a) shows, Cleft sentences involve a copula at the end of the sentence. The copula is optional in the case of embedded Cleft, as shown in (7b). This fact explains the optionality of the copula in JS under the Cleft analysis.

(7) a. [TopP [Mary-ga kat-ta no]-wa [FocP ringo-o da]].
   Mary-Nom buy-Past C-Top apple-Acc be
   ‘It is an apple that Mary bought.’

b. Mary-ga nanika-o kat-ta ga, boku-wa
   Mary-Nom something-Acc buy-Past but I-Top
   Mary-Nom buy-Past C-Nom what-Acc (be) Q know-not
   ‘Mary bought something, but I don’t know what it is that Mary bought.’

On the other hand, regular wh-questions do not involve a copula as shown in (8), which would be a problem for the wh-movement analysis.

(8) Mary-ga nanika-o kat-ta ga, boku-wa
   Mary-Nom something-Acc buy-Past but I-Top
   [CP (sore-ga) nani-o (da) ka] wakara-nai.
   (it-Nom) what-Acc (be) Q know-not
   ‘Mary bought something, but I don’t know what Mary bought.’

Second, JS optionally has a pronominal subject *sore-ga* ‘it-Nom’ before the wh-remnant as shown in (9).

(9) Mary-ga nanika-o kat-ta ga, boku-wa
   Mary-Nom something-Acc buy-Past but I-Top
   [CP (sore-ga) nani-o (da) ka] wakara-nai.
   (it-Nom) what-Acc (be) Q know-not
   ‘Mary bought something, but I don’t know what it is.’

---

\(^2\) Another argument against the wh-movement analysis is that, as Kuwabara (1997) points out, JS is possible with a non-wh remnant (e.g. (i)), unlike English Sluicing (see also Takahashi 1994). Note that (i) could also be classified as embedded Stripping as it involves a non-wh remnant (See footnote 1).

(i) Haha-wa [CP [IP boku-no rusutyuu-ni Tanaka-ga tazuneteki-ta] to] Mother-Top I-Gen absence-during Tanaka-Nom come-to-see-Past that it-ta ga, boku-wa Tanaka-ga to omowa-nai. say-Past but I-Top Tanaka-Nom that think-not
   ‘Mother said that Tanaka came to see me during my absence, but I don’t think that Tanaka.’
This fact can be accommodated under the Cleft analysis, if the pronominal subject is somehow derived from the clausal subject of Cleft. A Cleft sentence has a presuppositional clause as its subject (the underlined part in (10)). We assume that there is an operation that replaces the presuppositional clause with a pronoun and thus derives the pronominal subject in (9).

(10) Mary-ga nanika-o kat-ta ga, boku-wa
  Mary-Nom something-Acc buy-Past but I-Top
  ‘Mary bought something, but I don’t know what it is that Mary bought.’

There is no place for a pronominal subject in a regular wh-question such as (8). So the fact in (9) is hard to capture under the wh-movement analysis.

Third, the wh-remnant in JS may appear with or without the Case-marker. For example, the Accusative Case-marker –o is optional in (11).

(11) Mary-ga nanika-o kat-ta ga, boku-wa
  Mary-Nom something-Acc buy-Past but I-Top
  [CP (it-Nom) what(-Acc) (be) Q know-not
  ‘Mary bought something, but I don’t know what it is.’

Note that the focus of a Cleft sentence also allows for Case-marker drop, as shown in (12a) (an example of matrix Cleft) and (12b) (an example of embedded Cleft).

(12) a. [TopP [Mary-ga kat-ta no]-wa [FocP ringo(-o) da]].
    Mary-Nom buy-Past C-Top apple(-Acc) be
    ‘It is an apple that Mary bought.’

b. Mary-ga nanika-o kat-ta ga, boku-wa
   Mary-Nom something-Acc buy-Past but I-Top
   ‘Mary bought something, but I don’t know what it is that Mary bought.’

Case-marker drop, on the other hand, is not allowed in regular wh-questions as shown in (13). Again, this is problematic for the wh-movement analysis of JS.

(13) Mary-ga nanika-o kat-ta ga, boku-wa
  Mary-Nom something-Acc buy-Past but I-Top
  [CP nani*(o) [IP Mary-ga ti kat-ta ka]] wakara-nai.
  what*(Acc) Mary-Nom buy-Past Q know-not
  ‘Mary bought something, but I don’t know what Mary bought.’

Fourth, Kizu (1997) shows that, in JS, a numeral quantifier must follow the sluiced wh-phrase it modifies. In the example in (14), the numeral quantifier takusan ‘many’ cannot precede the modified wh-phrase nani-o ‘what-Acc’.

- 202 -

(14) Mary-ga nanika-o takusan kat-ta ga, boku-wa
   Mary-Nom something-Acc many buy-Past but I-Top
   {[nani-o takusan/*takusan nani-o] ka] wakara-nai.
   what-Acc many/ many what-Acc Q know-not
   ‘Mary bought something a lot, but I don’t know what a lot.’

Kizu shows that this restriction also applies to Cleft, as shown in (15a) (matrix Cleft) and (15b) (embedded Cleft), supporting the Cleft analysis of JS.

(15) a. [Mary-ga kat-ta no]-wa {ringo-o takusan/*takusan ringo-o} da.
    Mary-Nom buy-Past C-Top {apple-Acc many/ *many apple-Acc} be
    ‘It is apples a lot that Mary bought.’

    b. Mary-ga nanika-o takusan kat-ta ga, boku-wa [CP [Mary-ga kat-ta
       Mary-Nom buy-Past but I-Top Mary-Nom buy-Past
       no]-ga {nani-o takusan/*takusan nani-o} ka] wakara-nai.
       C-Nom what-Acc many/ *many what-Acc Q know-not
       ‘Mary bought something a lot, but I don’t know what a lot it is that Mary bought.’

There is no such restriction on wh-questions in general; the quantifier can either precede or follow the wh-phrase in (16).

(16) Mary-ga nanika-o takusan kat-ta ga, boku-wa
    Mary-Nom something-Acc many buy-Past but I-Top
    {nani-o takusan/takusan nani-o} Mary-ga kat-ta ka wakara-nai.
    what-Acc many/ many what-Acc Mary-Nom buy-Past C know-not
    ‘Mary bought something a lot, but I don’t know what a lot Mary bought.’

Finally, the island restrictions of JS are the same as that of Cleft. We have seen in (11) and (12) that JS and Cleft may or may not have a Case-marker on its remnant/focus phrase. But as Hoji (1990) shows, Case-marked and non-Case-marked versions have different status with respect to island sensitivity. As shown in (17), Case-marked JS is island sensitive, while non-Case-marked JS is not.

(17) Mary-ga [NP [CP aru ronbun-o kai-ta] hito]-o hihansi-ta ga,
    Mary-nom some paper-acc write-Past person-acc criticize-Past but
    boku-wa dono ronbun(*-o) ka wakara-nai
    I-Top which paper(*-Acc) Q know-not
    ‘Mary criticized [a person who wrote some paper], but I don’t know which paper.’

The same Case-marked/non-Case-marked distinction exists in Cleft; (18) shows that Case-marked Cleft, but not non-Case-marked Cleft, is island-sensitive.

(18) [CP Mary-ga [NP [CP kaita] hito]-o hihansi-ta no]-wa
     Mary-nom write person-acc criticize-Past C-top
     kono ronbun(*-o) da.
     this paper(*-Acc) be
     ‘Lit. Lit. :It was this paper that Mary criticized [the person who wrote t1].’

This fact follows if Cleft and JS share the same derivation, while such a similarity is not accounted for under the wh-movement analysis.

Based on these previous observations, we adopt the Cleft analysis of JS.
2.3. Arguments for the Cleft Analysis of JWS

This subsection shows that the arguments for the Cleft analysis reviewed above hold for JWS as well, supporting the Cleft analysis of JWS.

First, JWS allows an optional copula *da ‘be’* in the same way as JS, as shown in (19).

(19) Mary-ga ringo-o kat-ta ga, boku-wa
    Mary-Nom apple-Acc buy-Past but I-Top
    [naze ringo-o (da) ka] wakara-nai.
    why apple-Acc (be) Q know-not
    ‘Mary bought an apple, but I don’t know why an apple.’

Second, JWS allows an optional pronominal subject *sore-ga ‘it-Nom’*. Note that the pronominal subject can either precede or follow the wh-phrase *naze ‘why’*, as exemplified in (20a) and (20b), respectively. We will come back to this observation in Section 4.

(20) Mary-ga ringo-o kat-ta ga, boku-wa …
    Mary-Nom apple-Acc buy-Past but I-Top …
       (it-Nom) why apple-Acc (be) Q know-not
       ‘Mary bought an apple, but I don’t know why it is an apple.’
    b. [naze (sore-ga) ringo-o (da) ka] wakara-nai.
       why (it-Nom) apple-Acc (be) Q know-not
       ‘Mary bought an apple, but I don’t know why it is an apple.’

Third, JWS allows Case-marker drop on the non-wh remnant; the non-wh remnant *ringo ‘apple’* in (21) may or may not have the Accusative Case-marker.

(21) Mary-ga ringo-o kat-ta ga, boku-wa
    Mary-Nom apple-Acc buy-Past but I-Top
    [(sore-ga) ringo-o  naze (da) ka] wakara-nai.
    (it-Nom) why apple-(Acc) (be) Q know-not
    ‘Mary bought an apple, but I don’t know why it is an apple.’

Fourth, there is a word order restriction between *naze ‘why’* and the non-wh remnant in JWS, a similar type of restriction to Kizu’s (1997) word order restriction between a numeral quantifier and the wh-remnant of JS. As exemplified in (22), the non-wh remnant (e.g. *ringo-o ‘Apple-Acc’*) must not precede the wh-phrase *naze ‘why’* in JWS.

(22) *Mary-ga ringo-o kat-ta ga, boku-wa
    Mary-Nom apple-Acc buy-Past but I-Top
    [(sore-ga) ringo-o  naze (da) ka] wakara-nai.
    (it-Nom) apple-Acc why (be) Q know-not

Crucially, the same restriction applies to Cleft with *naze ‘why’* and a non-wh remnant, as shown in (23) (Kawamura 2007); when both *naze* and a non-wh phrase reside in the Cleft focus position, *naze* must precede the non-wh phrase.

(23) [Mary-ga kat-ta no]-wa [?naze ringo-o/ *ringo-o naze] na no?
    Mary-Nom buy-Past C-Top why apple-Acc/ *apple-Acc why be Q?
    ‘Why is it an apple that Mary bought?’
On the other hand, there is no such restriction on wh-questions with *naze ‘why’, as is indicated by the fact that *naze can precede or follow *ringe-o ‘apple-Acc’ in (24).

(24) {Naze ring-o-o/ Ringo-o naze} Mary-ga kat-ta no?
    Why apple-Acc/ apple-Acc why Mary-Nom buy-Past Q
    ‘Why did Mary buy an apple?’

Finally, there is a Case-marked/non-Case-marked distinction in JWS with respect to island phenomena in the same way as Cleft and JS; Case-marked JWS is island-sensitive, while non-Case-marked JWS is not.

(25) Mary-wa [NP [|CP suruusing-no ronbun-o kaita] hito] -o hihansi-ta ga,
    Mary-top Sluicing-gen paper-acc wrote person-acc criticize-Past but
    boku-wa naze suruusing-no ronbun(*-o) (da) ka wakara-nai.
    I-top why Sluicing-gen paper (-acc) (be) Q know-not
    ‘I heard Mary criticized a student who wrote a paper on Sluicing, but I don’t know why a paper on Sluicing.’ (asking for the reason of criticizing)

These five arguments all indicate that JWS, as well as JS, has properties similar to Cleft, and supports the analysis where JWS is derived in the same way as Cleft, rather than as regular wh-questions.

3. JWS is not Multiple Cleft

As we have mentioned in the introduction, JWS is apparently similar to Multiple Sluicing such as (3) in that there are multiple remnants. JWS involves a wh-remnant *naze ‘why’ and a non-wh remnant, while Multiple Sluicing involves two wh-remnants. 3 However, these constructions behave differently with respect to Case-marker drop.

Hiraiwa and Ishihara (2002), who claim that JS is derived in the same way as Cleft, claims that Multiple Sluicing and Multiple Cleft share the same derivation. Their argument is based on the fact that both Multiple Sluicing and Multiple Cleft resist Case-marker drop, as exemplified in (26) and (27), respectively.

(26) Dareka-ga nanika-o kat-ta ga, boku- wa
    Someone-Nom something-Acc buy-Past but I-Top
    [dare*(ga) nani*(o) ka] wakara-nai.
    who*(Nom) what*(-Acc) Q know-not
    ‘Someone bought something, but I don’t know who what.’

(27) [Kat-ta no]-wa Mary*(ga) ringo*(o) da.
    Buy-Past C-Top Mary*(Nom) apple*(Acc) be
    ‘It was Mary, an apple that e1 bought e2.’

Kawamura (2007), on the other hand, discusses Cleft with *naze and a non-wh focus (e.g. (23)) and shows that such a sentence allows Case marker drop, as shown in (28).

---

3 As JS is possible with a non-wh remnant (See footnote 2), the fact that the second remnant is non-wh in JWS is not a priori reason to distinguish it from Multiple Sluicing.
Thus she concludes that sentences such as (23)/(28) should be distinct from Multiple Cleft.

We can conclude the same for the JWS, as JWS allows Case-marker drop as we have already seen in (21). JWS should not be treated merely as Sluicing which has multiple remnants. Especially, it will be argued in Section 4 that the wh-remnant naze ‘why’ in JWS has a special status distinct from Sluicing remnants in JS. Note that Sluicing with a wh-phrase other than naze ‘why’ and a non-wh remnant resists Case-marker drop, as shown in (29).

(29) Minna-ga nanika-o potluck-ni motteki-ta ga, boku-wa
Everyone-Nom something-Acc potluck-to bring-Past but I-Top
[dare*(-ga) pan*(-o) ka] wakara-nai.
who*(-Nom) bread*(-Acc) Q know-not
‘Everyone brought something to the potluck, but I don’t know who bread.’

That is, such a sentence has a property of Multiple Sluicing. Only JWS, which involves naze ‘why’ should be differentiated from these types of sentences.

Interestingly, JWS may have more than one non-wh remnant in addition to naze ‘why’, as shown in (30). In this case, the multiple non-wh remnants disallow Case-marker drop, indicating that this is a case of Multiple Sluicing.

(30) John-ga pan-o potluck-ni motteki-ta ga, boku-wa
John-Nom bread-Acc potluck-to bring-Past but I-Top
[naze John*(-ga) pan*(-o) ka] wakara-nai.
why John*(-Nom) bread*(-Acc) Q know-not
‘John brought bread to the potluck, but I don’t know why John bread.’

In order to reveal the special status of naze ‘why’ in ellipsis, we will discuss the source of the difference between ‘why’ and other wh-phrases in Section 4.

4. Base Generation of ‘Why’ and Focus

4.1. Special Properties of ‘Why’

The observation that why has different properties from other wh-phrases is not new in the literature. For example, Ko (2005) argues that Korean/Japanese ‘why’ is base-generated in [Spec, CP], while other wh-phrases in these languages undergo covert wh-movement to [Spect, CP] at LF.

Her argument is based on data of intervention effects between wh-phrases and Negative Polarity Items (NPIs) such as –sika ‘only’. (31a) shows that there is an intervention effect when a wh-phrase is lower than ‘only’. On the other hand, there is no intervention effect when ‘why’ is lower than ‘only’ as in (31b).

(31) a. *Hanako-sika nani-o yoma-nakat- no?
Hanako-only what-Acc read-not-Past Q
‘What did only Hanako read?’
b. Hanako-sika naze sore-o yoma-nakat- no?
Hanako-only why it-Acc read-not-Past Q
‘Why did only Hanako read it?’
She attributes the lack of intervention effect with ‘why’ and the NPI ‘only’ to the fact that, unlike other wh-phrases, ‘why’ is base-generated in [Spec, CP]. (32a) and (32b) illustrate the structure of (31a) and (31b), respectively. In (32a), the object wh-phrase ‘what’ undergoes LF wh-movement beyond the NPI subject, which causes an intervention effect. ‘Why’ in (32b), on the other hand, is base-generated in [Spec, CP] and thus does not need to undergo LF-movement. The NPI subject ‘Hanako-only’ in this example has undergone scrambling past ‘why’ but this movement, under Ko’s assumption, does not induce intervention effects, unlike LF wh-movement. This is why only wh-phrases other than ‘why’ exhibit the intervention effects.4

(32) a. *[[CP [IP Hanako-only] what-Acc read-not-Past] Q]  
    *intervention with LF-movement

    b. Hanako-only [CP why [IP t1 it-Acc read-not-Past] Q]  
    no intervention with scrambling

Based on her argument, we will assume below that naze ‘why’ in JWS is base-generated in [Spec, CP].

Then, what is the status of the non-wh remnant in JWS? We claim that the non-wh remnant is a focus phrase that is ‘associated’ with why. Bromberger (1992) observes that why, unlike other wh-phrases, can induce the so-called ‘association with focus’ effect. For example, the answer to a why-question changes depending on focus, as the two examples in (33) shows. When the subject Adam is focussed and thus is associated with why, the question is asking why Adam but not someone else ate the apple, while when the object the apple is focused, the sentence is asking for the reason why Adam ate the apple instead of some other food.

(33) a. A: Why did ADAM eat the apple?  
    B: Because he (Adam) is the one that Eve worked on.

    b. A: Why did Adam eat the APPLE?  
    B: Because it (the apple) was the only food around.

No such effect, on the other hand, is observed with wh-phrases other than why.

(34) a. A: When did ADAM eat the apple?  
    B: At 4 p.m. on July 7, 24,000,000 B.C.

    b. A: When did Adam eat the APPLE?  
    B: At 4 p.m. on July 7, 24,000,000 B.C.

---

4 The lack of intervention effects with ‘why’ does not extend to ‘why’ inside a declarative clause (e.g. (i)). Ko (2005) suggests that such ‘why’ must undergo LF-movement.

    Hanako-only Taro-Nom why come that say-not-Past Q
    ‘What is the reason x such that only Hanako said that Taro will come for x?’

Naze in the JWS example (ii) apparently corresponds to the one in (i). (ii), however, cannot mean (iib). Only the Cleft interpretation in (iia) is possible. (For (ii) to mean (iib), the phrase Taro(-ga) must be absent.) Therefore, the type of naze in (i) is irrelevant to our JWS argument.

(ii) Hanako-ga [Taro-ga kuru to] it-ta ga, naze Taro(-ga) ka wakakara-nai.
    Hanako-NomTaro-Nom come that say-Past but why Taro(-Nom) Q know-not
    ‘Hanako said that Taro came, but I don’t know why Taro.’
    a. … I don’t know why it is Taro that Hanako said will come.
    b. *… I don’t know what is the reason x such that Hanako said that Taro will come for x.
We follow Bromberger in that only *why* has the ability to induce association with focus, and the associated focus shows up as the second remnant of JWS after *naze* ‘why’. This is why other wh-phrases do not have a construction parallel to JWS (as shown in the fact that (29) is merely treated as Multiple Sluicing).

4.2. The Structure of JWS

Our conclusions so far are: the underlying structure of JWS is Cleft, *naze* ‘why’ in JWS is base-generated in [Spec, CP], and the non-wh phrase in JWS is a focused phrase associated with ‘why’. This subsection proposes an analysis of JWS which is compatible with these conclusions. The analysis is based on Hiraiwa and Ishihara’s (2002) analysis of JS.

Hiraiwa and Ishihara (2002) claim that both Cleft and JS are derived from the ‘*no da*’ in-situ focus construction. The fact that the two constructions share the same underlying structure can account for the similarities between JS and Cleft, which are arguments for the Cleft analysis of JS in Section 2.

(35) illustrates the derivation of Cleft under Hiraiwa and Ishihara’s analysis. (35a) is an example of the ‘*no da*’ in-situ focus construction. They assume that *no* is the Fin head and *da* is the Foc head. When Cleft is derived from this structure, a focused phrase inside TP undergoes focus movement to [Spec, FocP] as illustrated in (35b), and then the rest of FinP undergoes remnant movement to [Spec, TopP] as shown in (35c), resulting in the word order of Cleft.

(35) a. $\left[ \text{FocP} \right]_{\text{Mary-ga}} \text{ringo-o kat-ta} \left[ \text{FinP} \left( \text{no} \right) \right] \left[ \text{Foc0 da} \right]$
   Mary-Nom apple-Acc buy-Past C be
   ‘It is that Mary bought an apple.’

   b. $\left[ \text{FocP} \right]_{\text{ringo-o t}} \left[ \text{FinP} \left( \text{Mary-ga} \right) \right] \left[ \text{TP} \right] \left[ \text{kat-ta} \right] \left[ \text{Fin0 no} \right] \left[ \text{Foc0 da} \right]$
   apple-Acc Mary-Nom buy-Past C be

   c. $\left[ \text{TopP} \right]_{\text{Mary-ga t}} \left[ \text{FinP} \left( \text{no} \right) \right] \left[ \text{TP} \right] \left[ \text{kat-ta} \right] \left[ \text{Fin0 no} \right] \left[ \text{Foc0 da} \right]$
   apple-Acc Mary-Nom buy-Past C-Top apple-Acc be
   ‘It is an apple that Mary bought.’

In order to capture the similarities between Cleft and JS reviewed in Section 2.2, Hiraiwa and Ishihara argue that JS is also derived from the ‘*no da*’ in-situ focus construction. The derivation is shown in (36). Starting from the ‘*no da*’ sentence in (36a), the wh-phrase undergoes focus movement to [Spec, FocP] as shown in (36b), in the same way as the focus movement in the Cleft derivation (35b). Instead of the remnant movement in (35c), FinP-deletion occurs in JS as shown in (36c). Thus, the similarity between Cleft and JS is attributed to the shared derivation in their analysis.

5 In this construction, any phrase inside the sentence followed by ‘*no da*’ gets a focused interpretation. Especially, if a phrase has a prosodic prominence, it is interpreted as the focus as shown in the contrast in (i).

(i) a. $\left[ \text{CP} \right]_{\text{TARO-ga}} \text{kono ringo-o tabe-ta no} \left[ \text{da} \right]$
   TARO-Nom this apple-Acc eat-Past C be
   ‘It is taro that ate this apple.’

   b. $\left[ \text{CP} \right]_{\text{Taro-ga}} \text{KONO RINGO-o tabe-ta no} \left[ \text{da} \right]$
   Taro-Nom THIS APPLE-Acc eat-Past C be
   ‘It is this apple that Taro ate.’

(36) Mary-ga nanika-o kat-ta ga, boku-wa
Mary-Nom something-Acc buy-Past but I-Top

a. [FocP [FinP [TP Mary-ga nani-o kat-ta] [Fin0 no]] [Foc0 da]] ka wakara-nai
Mary-Nom what-Acc buy-Past C be Q know-not

b. [FocP nani-o [FinP [TP Mary-ga t1 kat-ta] [Fin0 no]] [Foc0 da]] ka wakara-nai
Mary-Nom what-Acc buy-Past C be Q know-not

c. [FocP nani-o [FinP [TP Mary-ga t1 kat-ta] [Fin0 no]] [Foc0 da]] ka wakara-nai
Mary-Nom what-Acc buy-Past C be Q know-not

‘Mary bought something, but I don’t know what (it is that Mary bought).’

Given that JWS also exhibits properties of Cleft (Section 2.3), we propose the FocP in JWS is derived in the same way as (36). The difference, however, is that the wh-phrase naze ‘why’ is base-generated in [Spec, CP] in the case of JWS, as shown in (37a) (See Section 4.1). The wh-phrase undergoes focus movement to [Spec, FocP] as shown in (37b). This movement is induced by the focus association effect of why. Finally, FinP-deletion occurs as shown in (37c) to derive the JWS sentence.

(37) Mary-ga ringo-o kat-ta ga, boku-wa …
Mary-Nom apple-Acc buy-Past but I-Top …

a. [CP naze [FocP [FinP [TP Mary-ga ringo-o kat-ta] [Fin0 no]] [Foc0 da]] ka
Mary-Nom apple-Acc buy-Past C be Q

b. [CP naze [FocP ringo-o [FinP [TP Mary-ga t1 kat-ta] [Fin0 no]] [Foc0 da]] ka
Mary-Nom apple-Acc buy-Past C be Q

c. [CP naze [FocP ringo-o [FinP [TP Mary-ga t1 kat-ta] [Fin0 no]] [Foc0 da]] ka
Mary-Nom apple-Acc buy-Past C be Q

… wakara-nai.
… know-not

‘Mary bought an apple, but I don’t know why (it is) an apple (that Mary bought).’

Recall that JWS may have a pronominal subject sore-ga ‘it-Nom’ as shown in (20), and we assumed that the pronominal subject should be analogous to the presuppositional clause of the Cleft. Under Hiraiwa and Ishihara’s (2002) analysis, it means that it should be analogous to FinP in (35c). We propose that, instead of FinP deletion in (37c), FinP may undergo topicalization in the same way as the Cleft example (35c), deriving the structure in (38).

(38) Mary-ga ringo-o kat-ta ga, boku-wa
Mary-Nom apple-Acc buy-Past but I-Top

[CP naze [TopP [FinP [TP Mary-ga t1 kat-ta] [Fin0 no]] ga [FocP ringo-o [FinP why Mary-Nom buy-Past C-Nom apple-Acc
gaga] [Foc0 da]] Top0]] wakara-nai
Mary-Nom buy-Past C be Q know-not

Then the topicalized FinP may be replaced with the pronominal subject, deriving the
sentences in (20b), repeated below as (39b). In the case of (39a), which is repeated from (20a), the pronominal subject is further scrambled above *nage ‘why’. This is an example of subject scrambling in the same manner as (32b) in Ko’s (2005) analysis.

(39) Mary-ga ringo-o kat-ta ga, boku-wa …
     Mary-Nom apple-Acc buy-Past but I-Top …
        (it-Nom) why apple-Acc (be) Q know-not
     b. [nage (sore-ga)] ringo-o (da) ka] wakara-nai.
        why (it-Nom) apple-Acc (be) Q know-not
‘Mary bought an apple, but I don’t know why it is an apple.’

In sum, our analysis based on Hiraiwa and Ishihara’s analysis can capture the similarities between Cleft, JS, and JWS while maintaining the claim by Ko (2005) that *why is base-generated in [Spec, CP] and Bromberger’s (1992) intuition that *why can be associated with a focused phrase.

4.3. Against an Alternative Analysis

Finally, let us give an argument against a possible alternative to the analysis in Section 4.2. Kawamura (2007) discusses Cleft with *nage ‘why’ and a non-wh focus, which we have seen share similar properties with JWS (i.e. the word order restriction in (23) and the possibility of Case-marker drop in (28)), and gives the following analysis. She claims that *nage ‘why’ is bound by a wh-operator in [Spec, CP], and that *nage and the non-wh focus in Cleft makes a constituent as shown in (40). There, *nage is the specifier and the non-wh focus is the complement of the same projection FocP.

(40) FocP
     nage why
     Foc’ ringo-o apple-Acc

One might take this analysis of *why-Cleft and extend it to the analysis of *nage and the non-wh remnant in JWS.

A problem with this analysis, however, is that more than one focused phrase can follow *nage ‘why’ in Cleft as shown in (41), as well as JWS (as we have already seen in (30)).

(41) [Potluck-ni motteki- ta no]-wa, nage John*(-ga) pan*(-o) na no?
     Potluck-to bring-Past C-Top why John*(-Nom) bread*(-Acc) be Q
     ‘Why is it John1 bread2 that e1 brought e2 to the potluck?’

---

6 The pronominal subject in JWS (e.g. (20)) has the Nominative marker –ga instead of the topic marker –wa in (38). Although we do not have a clear account of this fact, it should be attributed to the general property in Japanese that embedded subjects tend to be marked by the Nominative marker rather than the topic marker. A presuppositional clause subject in Cleft, too, is more natural when it is marked by –ga as shown in (10).
To accommodate (41), Kawamura must assume a multiple-complement structure, which is not a standard assumption. Moreover, (41) is not asking for the reason “why it is John” and the reason “why it is bread” simultaneously. Rather, it is asking “why there holds a bring-relation between John and bread”. Under this interpretation, ‘John’ and ‘bread’ shouldn’t be merely coordinated in the one and the same complement position of FocP, either. It would be hard to accommodate such examples under the assumption where naze and the non-wh phrase(s) make a constituent.

5. Conclusion

In this paper, we have seen that Japanese Why-Stripping (JWS) has similar properties with Japanese Sluicing (JS) and Cleft, supporting the Cleft-based analysis of both JS and JWS. Although JWS involves both naze ‘why’ and (a) non-wh remnant(s), it shows a distinct property from multiple Cleft (i.e. the possibility of Case-marker drop). Given these data, we give JWS the structure illustrated in (37c). We assume that JS and JWS are both derived from the ‘no da’ in-situ focus construction (Hiraiwa and Ishihara 2002). JWS, however, is not a mere example of Cleft in that naze ‘why’ is base-generated in [Spec, CP] and the only focus of the Cleft is the non-wh remnant which resides in [Spec, FocP]. It is associated with why at LF, which is a special property of why according to Bromberger (1992). According to this view, naze ‘why’ in JWS is base-generated in the left-periphery, and the non wh remnant undergoes focus movement. Evidence has been provided that JWS is distinct from Multiple Sluicing.

References


FUSION AND SCATTERING IN PARTICLE COLLISIONS:
WHEN UNIVERSAL QUANTIFICATION AND CONTRASTIVE FOCUS INTERACT
WITH THEMSELVES AND EACH OTHER IN VIETNAMESE

Chieu Nguyen
University of Chicago

1. Introduction

Universal quantification may be expressed in Vietnamese\(^1\) by the combination of
wh-morphology and a preverbal particle cũng, following a paradigm that also licenses
interrogative, existential, and negative polarity interpretations for wh-indeterminates in
different licensing contexts. The universal construction is peculiar in that it requires the
licensed wh-phrase to precede its apparent licensor cũng in the clause.\(^2\) Given the basic SVO
constituent order of Vietnamese, this results in the obligatory fronting of a universally
quantified object (1), whereas the other wh-indeterminate constructions allow an object to
remain in situ.\(^3\)

(1) a. * [ Con quái vật ] cùng phá hoại [ thành phố nào ].
          CL monster CUNG destroy city which

b. [ Thành phố nào ] [ con quái vật ] cùng phá hoại.
          city which CL monster CUNG destroy

‘The monster destroyed every city.’

If cùng is to be regarded as a licensing operator, as in Tran (2009), the lack of adjacency
between it and the preposed wh-indeterminate in (1b) complicates a straightforward
compositional account of the semantics of the construction, placing the licensee outside the
scope of the licensor. In addition, the possibility of multiple quantifications in (2), a property
not found in similar wh-constructions in Japanese and Mandarin in the same manner, presents
a further confound where the function of multiple operators is expressed by what appears to
be a single licensor, a phenomenon I will refer to as fusion.\(^4\)

\(^*\) I would like to thank Karlos Arregi, Rebekah Baglini, Karen Chu, Mitch Erlewine, Itamar
Francez, Anastasia Giannakidou, James Huang, Yoshihisa Kitagawa, Peter Klecha, Greg Kobele,
Jackson Lee, Jason Merchant, Nga Nguyen, Lien Pham, Mike Pham (no relation), Diane Rak, Osamu
Sawada, Sanae Tamura, and Masahiro Yamada, and audiences in Chicago, Tsu, and Boston for their
advice, feedback, and judgments. All errors are my own.

\(^1\) The data come from a variety of Southern Vietnamese which differs in some significant ways from
other varieties discussed in the literature. Even more significantly, my own variant of the variety
shows some idiosyncratic properties; other speakers reject some of the more complex sentences I offer
in support of my analysis. Differences are noted where present, and I will attempt to account for them.

\(^2\) There are a few other constructions that produce universal interpretations, such as the prenominal
mỗi ‘every’ or tất cả ‘all’, but they do not involve wh-morphology and do not require fronting.

\(^3\) Glossing abbreviations are as follows: 3SG = third person singular, ACC = accusative, CL =
classifier, DEM = demonstrative, DEO = deontic, FUT = future, IPFV = imperfective, NEG = negation,
NOM = nominative, PFV = perfective, PL = plural, POSS = possessive, PST = past, Q = question particle,
TOP = topic, VRF = verum focus.
To account for this problem, I argue that cùng itself is not the licensing operator, a role fulfilled by an unpronounced focal head adjacent to the fronted phrase, but is instead a pronoun generated within vP that is extracted to the edge of TP. This requires there to be an instance of cùng for each focal operator, but only one instance is pronounced, so the fusion effect remains to be explained, but it is at least localized to the edge of TP rather than extending across multiple heads in the left periphery.

Another interesting phenomenon is the interaction between the universal quantification construction and a syntactically parallel contrastive focus construction in which the particle mới plays a role similar to cùng. The mới construction requires preposing the phrase containing an associated focused element (3a) and allows licensing focused elements in separate constituents (3b). It can also cooccur with universal quantification in the same clause, resulting in both particles being pronounced.6

In these cases, the particles do not appear in a universal order, but they are also not in free variation: in some situations they surface as cùng mới, and in others, mới cùng. The order is determined not by the relative scope of the operators, which matches the order of the fronted constituents, but by the positions of the verbal arguments they correspond to: the default order is cùng followed by mới, unless an external argument receives contrastive focus, in which case mới appears before cùng. This conditioned variation I refer to as scattering.

In this paper, I propose that the particles cùng and mới are syntactically distinct from the focal heads that license their associated interpretations, and that the fusion and scattering properties of the two constructions when more than one focal head appears in a clause follow from a collision of pronominal particles at the edge of TP, resulting in a fusion effect through haplology applied to adjacent instances of the same particle and a scattering effect triggered by movement of external arguments to check agentive Case.
2. Universal Quantification

Vietnamese follows a typologically common pattern in which wh-phrases can produce a number of interpretations under different licensing contexts (Haspelmath 1997); interrogative, existential, universal, and negative polarity ‘any’ interpretations are all possible (4). Only in the universal construction is it required that the wh-indeterminate be preposed (4c).

(4) a. Interrogative
Con quái vật phá hoại [ thành phố nào ] vậy?
CL monster destroy city which Q
‘Which city did the monster destroy?’

b. Existential
Con quái vật có phá hoại [ thành phố nào ].
CL monster VRF destroy city which
‘The monster did destroy some city.’

c. Universal
[ Thành phố nào ] con quái vật cũng phá hoại.
city which CL monster CUNG destroy
‘The monster destroyed every city.’

d. Negative polarity ‘any’
Con quái vật không phá hoại [ thành phố nào ].
CL monster NEG destroy city which
‘The monster did not destroy any city.’

In a universal construction, the preposed phrase is necessarily D-linked (Pesetsky 1987), requiring the domain of quantification to be accessible in the discourse. There is also a non-singleton constraint: the use of a wh-universal presupposes that the domain have a cardinality strictly greater than 1; (4c) is only felicitous if there are at least two cities in the domain.

There is also an additive use of cũng ‘also’/‘even’, shown in (5), but this does not require the associated phrase to be fronted. While it would be ideal to unify additive and universal cùng, their syntactic differences suggest that they might not be related synchronically, so I will set additive cũng aside as a distinct construction.

I also read CL book that already.
‘I too have read that book.’ / ‘I have read that book too.’

---

7 The conditions licensing these interpretations are somewhat complicated and are not consistent across dialects of Vietnamese; Bruening and Tran (2006) discuss the conditions licensing the interrogative, and Tran (2009) discusses the conditions licensing the existential. Free choice any is also found in wh-indeterminate paradigms in other languages (Giannakidou and Cheng 2006), but in Vietnamese, free choice constructions can be reduced to the combination of a universal and an external modal element.

8 There appears to be a weaker constraint against domains of size 2 that holds in some situations: (4c) is slightly odd, though not entirely infelicitous, if there are exactly two cities. This constraint appears not to hold if the domain consists of two alternatives; for example, if the speaker is deciding between two roads to take, it is perfectly acceptable to use the wh-universal to quantify over the domain of possible roads. In any case, a domain of size 3 is acceptable but not a domain of size 1.

9 For Tran (2009), it does require fronting; there seems to be dialectal variation here. There is also possibly a contribution from rời ‘already’, which if omitted lowers the acceptability of the object focus reading in (5a).
b. [ Cuốn sách đó ] [ tôi ] cúng đọc rồi.
   CL book that I also read already
   ‘That book, I too have read.’ (topicalized object) / ‘That book too I have read.’

A particle expressing both an additive meaning and universal quantification in conjunction with a wh-indeterminate can be found in other languages, including Japanese, in which a universal reading arises from the combination of a particle -mo and a wh-indeterminate such as dono ‘which’ inside the phrase immediately preceding it (Nishigauchi 1986, 1991; Shimoyama 2001; Yoshimura 2007), as in (6).¹⁰

(6) Japanese (Shimoyama 2001: 2)
   Yoko-wa dono hon-mo yonda.
   Yoko-TOP which book-MO read
   ‘Yoko read every book.’

A similar universal construction in Mandarin consists of a particle dōu and a preposed wh-indeterminate such as shéi ‘who’ or shénme ‘what’ to its left, as in (7), though not necessarily adjacently; a preposed wh-indeterminate may be separated from dōu by the clausal subject, as with cúng in Vietnamese, though it is not possible to do so with an adverb (Cheng 1991, 1995; Giannakidou and Cheng 2006).

(7) Mandarin (Cheng 1995: 202)
   a. Shéi dōu hui lái.
      who all will come
      ‘Everyone will come.’
   b. Zhāngsān shénme dōu chī.
      Zhangsan what all eat
      ‘Zhangsan eats everything.’

The constructions in all three languages allow the long-distance binding of an embedded wh-phrase: the wh-phrase need not be the constituent adjacent to the apparent licensor (8).¹¹

(8) a. Japanese (Shimoyama 2001: 2)
   which book-ACC read child -MO well slept
   ‘For every book x, the child who read x slept well.’

¹⁰ This should be distinguished from NPI -mo (ia), which Yoshimura (2007) observes to have a distinct pitch pattern. Likewise, Vietnamese cúng is not used in NPI wh-indeterminate constructions (ib), and the use of cúng with negation (ii) is just a universal with a negated predicate and requires a D-linked wh-indeterminate (ai ‘who’ has both D-linked and non-D-linked uses).

   John-TOP who-MO see-NEG-PST John NEG see who
   ‘John didn’t see anyone.’

(ii) Vietnamese cúng with negation (not NPI)
   Ai John cúng không thấy.
   who John CUNG NEG see
   ‘John didn’t see any of the people.’

¹¹ Some Vietnamese speakers report a difficulty in accepting sentences in which a wh-indeterminate is embedded in a relative clause modifying a non-wh-indeterminate, as in (8b). Other embedded wh-indeterminates, as in (8d), are acceptable.
Furthermore, all three constructions also allow same-constituent licensing: two or more wh-phrases can be bound within a constituent with only one apparent licensor (9).

\[(9)\]
\[
a. \text{Japanese (Shimoyama 2001: 26)}
\]
\[
[ [ [ \text{Dono gakusei-ga} \text{ [ [ dono ie-ni ] syootaisita ] pianisuto ] -mo kita.} \\
\text{which student-NOM which house-to invited pianist -MO came}
\]
\text{‘Every pianist that a student had invited to some house came.’}
\]

b. \text{Vietnamese}

\[
[ \text{Nhạc sĩ pianô [ [ học sinh nào ] mời tới [ nhà nào ]] cúng bò chạy.} \\
\text{Musician piano student which invitecome house which cung flee}
\]
\text{‘Every pianist that some student had invited to some house fled.’}
\]

c. \text{Mandarin (Cheng 1995: 222)}

\[
[ \text{Shéi chí shènme } \text{ dōu gēn wǒ wúguān.} \\
\text{who eat what all to I irrelevant}
\]
\text{‘Whoever eats whatever is irrelevant to me.’}
\]

d. \text{Vietnamese}

\[
[ \text{Ai ăn cái gì} \text{ cúng phải đi nhà thương.} \\
\text{who eat what CUNG must go hospital}
\]
\text{‘Whoever ate anything had to go to the hospital.’}
\]

The Vietnamese construction differs in that it also allows separate-constituent licensing: quantification over wh-phrases in two distinct constituents that are not themselves part of a larger constituent external to the licensor (10).\[12\]

\[(10)\]
\[
a. \text{Japanese (Shimoyama 2001: 26)}
\]
\[
[ [ Món nào ] \text{ [ ai ăn nó ] cúng chê.} \\
\text{dish which who eat it CUNG antipraise}
\]
\text{‘For every dish, everyone who ate it expressed a low opinion of it.’}
\]

b. \text{Vietnamese}

\[
[ \text{Ở đâu } [ \text{ ai } ] \text{ cúng muốn hòa bình.} \\
\text{at where who CUNG want peace}
\]
\text{‘Everywhere, everyone wants peace.’}
\]

This is in contrast to Mandarin, in which multiple separate wh-indeterminates cannot license multiple universal quantifications. Only the wh-indeterminate closest to dōu can be interpreted as a universal, and the others must receive interrogative interpretations (11).

\[12\] For Tran (2009), only the wh-phrase closest to cúng can attain universal quantification, and the others become interrogatives, as in Mandarin (11). Speakers of the present dialect, however, do accept wh-phrases in separate constituents.
In Japanese, multiple separate wh-indeterminates can be licensed, but only by multiple instances of -mo (12a). If the Vietnamese construction is structurally analogous, we would expect multiple instances of cúng as well. However, this is not the case (12b).

I will argue that (12a) and (12b) are nevertheless analogous structures, and that the apparent discrepancy is simply due to -mo and cúng not playing the same syntactic role. Instead, Japanese -mo corresponds to an unpronounced operator in Vietnamese, while cúng is a pronoun that moves to the edge of TP; its Japanese counterpart, if any, is unpronounced.

3. Same-Constituent vs. Separate-Constituent Licensing

At first glance, the same-constituent and separate-constituent licensing constructions in Vietnamese and Japanese might appear to have the same semantics, where the denotations given for (13a) and (13b) are logically equivalent.

Structurally, however, (13b) consists of two quantifications, an outer one and an inner one within its scope; the inner quantification has a domain dependent on particular assignments to the outer quantification variable. (13a) has only one quantification over child-book pairs, following unselective binding accounts of multiple wh-indeterminates licensed by a single operator, such as Kratzer and Shimoyama’s (2002) for Japanese, that allow a single quantifier to bind tuples of variables rather than a single variable (Lewis 1975). I argue that unselective binding does not reflect the structure of (13b).

To distinguish the two, we can examine the constraints that hold for any universal quantification. Both Vietnamese and Japanese have a constraint against singleton domains, which we can observe with the sentences in (14).
When the domain consists of three children, and each of them laughs, both (14a) and (14b) are acceptable, but when the domain consists only of one child, neither is acceptable (Osamu Sawada, p.c.).

What happens when we apply this test to the multiple wh-indeterminate examples? Here, if we have a context in which there are three children and three books, and each child reads all three of the books and praises all of them, then the same-constituent and separate-constituent licensing examples should both be felicitous. (13a) quantifies over 9 child-book pairs, while (13b) quantifies first over 3 children, and then for each child, the 3 books that that child read. In both cases, all quantifications have a domain size of at least 3. Indeed, the prediction holds.

When we change the context to have only one book per child, however, we expect a difference in acceptability resulting from the non-singleton constraint not being satisfied for the inner quantification. If Alice reads only the Aeneid and praises it, Bob reads only Beowulf and praises it, and Carol reads only The Thackery T. Lambshead Pocket Guide to Eccentric & Discredited Diseases and praises it, then (13a) will quantify over 3 child-book pairs (⟨Alice, Aeneid⟩, ⟨Bob, Beowulf⟩, ⟨Carol, TTTLPGiE&DD⟩). (13b), however, will first quantify over 3 children, and then for each assignment of the variable of quantification to a particular child, quantify over the set of books that that child read: for each child, the inner quantification has a singleton domain, and the non-singleton constraint fails to hold. This correctly predicts the observation that in this context, (13a) is felicitous, but (13b) is not.

If the Vietnamese separate-constituent licensing construction is structurally analogous to the Japanese construction with multiple instances of -mo, then we would predict the Japanese counterparts to the Vietnamese sentences, given in (15), to have the same cardinality-based felicity conditions.

(15) a. Same-constituent licensing


which book-ACC read which child -MO DEM-ACC praised

‘Every child who read any book praised it.’

∀(x, y) [CHILD(x) ∧ BOOK(y) ∧ READ(y)(x) ⇒ PRAISE(y)(x)]

b. Separate-constituent licensing


which child -MO DEM-NOM read which book -MO praised

‘Every child praised every book s/he read.’

∀x [CHILD(x) ⇒ [∀y:BOOK(y) ∧ READ(y)(x) ⇒ PRAISE(y)(x)]]

The prediction is borne out: when the three children each read all three books, both (15a) and

---

13 Each wh-indeterminate also carries its own non-singleton constraint: the context must have more than one individual satisfying the nominal description of each wh-indeterminate. This constraint parallels the referential vagueness property found for existential indeterminates (Giannakidou and Quer 2011) and may derive from properties common to all indeterminates. So (13a) is not felicitous in a context in which there is only one child and that child reads three books, because although there are three pairs, there is only one child. Likewise, it is not felicitous in a context in which there is only one book and three children read it. (13a) can be rendered more felicitous in these cases by adding extra children who read no books, or by adding extra books read by no children, respectively. This observation is consistent with non-singleton constraints on both the domain of universal quantification and the satisfiability set of each wh-indeterminate.
Thus we see that separate-constituent licensing in Vietnamese requires separate operators, but they are marked by only one apparent licensor, a fusion phenomenon that must be accounted for. The solution I will adopt is that Vietnamese has an unpronounced operator playing the role of -mo; this explains why there is no cũng next to the first fronted constituent and also allows for the separate-constituent licensing constructions in Vietnamese and Japanese to be handled in a uniform manner. Following Tran (2009), I characterize the universal quantification construction as having a Focus head with the wh-indeterminate in the Spec position, but treat this head as being unpronounced rather than assigning cùng to it.\footnote{Tran (2009) proposes that the wh-indeterminate moves to the Spec of FocP from within the vP. At least in the present dialect, there is evidence that the wh-indeterminate can be base-generated in Spec of FocP, though the situation is complicated by the existence of conditions that do show evidence of movement, for example when licensing a wh-indeterminate embedded within a pied-piped PP complement. When a DP is fronted, however, there is a lack of expected crossover effects (ia) associated with A’-movement (ib). The construction shares a number of syntactic properties with Clitic Left Dislocation in Romance and Greek (Cinque 1990; Iatridou 1995), though I set aside these details for now.}

The particle cùng is instead a pronoun that moves to the edge of TP; it refers to the variable of quantification introduced by a corresponding Focus head in the left periphery. Since this is a one-to-one relation, there must also be multiple instances of cùng in separate-constituent licensing constructions. I argue that these instances are adjacent and are reduced by a process of haplology, accounting for the fusion property.

An alternate solution would be to suppose that a single head has more than one specifier, allowing multiple wh-indeterminates to be licensed, in a fashion similar to some accounts of multiple fronted wh-interrogatives in languages like Bulgarian (Rudin 1988). To argue against this, I examine cases in which the syntactically parallel contrastive focus mới construction interacts with universal quantification. Since the same clause can host multiple instances of both constructions in an interspersed order, each operation taking scope over the next, a single operator head cannot handle all of the universal quantification operations in the clause.

4. Contrastive Focus

Contrastive focus may be expressed with the particle mới and a phrase to the left of it containing a focused element marked with prosodic stress; in the absence of stress on a particular element, the entire preposed phrase is focused. This is distinct from presentational focus, which is marked in situ by prosodic stress only and does not require the preposed position that contrastive focus with mới requires. This distinction in Vietnamese is consistent with É. Kiss’s (1998) separation of identificational focus and information focus, the former expressing exhaustive identification and occupying the specifier of a functional projection in the syntax, and the latter conveying new information without syntactic reordering.

The fronted phrase in a mới construction identifies a property which must hold for every individual satisfying the clausal predicate. When the fronted phrase is a definite DP, this identifies a unique individual (16a); in other circumstances, it can restrict the satisfiability set

   child which friend 3SG CUNG insult
   ‘Every child was insulted by his/her friend.’
   I NEG know child which friend 3SG insult
   ‘I don’t know which child was insulted by his/her friend.’
of the clausal predicate (16b).

(16) a. [Nhà tôi] nó mới tới.
   house I it come
   ‘It was my house that it came to.’

b. [Một con mèo rô bô màu xanh] mới xuất hiện.
   one cat robot color blue come
   ‘It was a blue robotic cat that appeared.’

The construction is syntactically parallel to universal cũng and similarly permits same-constituent and separate-constituent licensing (17).

(17) a. Same-constituent licensing
   child this read book which MOI praise it
   ‘It was this child-book pair such that the child praised the book.’

b. Separate-constituent licensing
   [Đứa này] [cuốn này nó đọc] mới __ khen __.
   child this book this 3SG read MOI praise
   ‘It was this child for whom it was this book that s/he read that s/he praised.’

The separate-constituent licensing case expresses two levels of contrast and can be used felicitously for instance when correcting a statement that is itself a correction of another statement.

As with universal quantification, there is a non-singleton domain constraint. Here it is the satisfiability set of the clausal predicate, rather than that of the preposed licensee, that is required to have a non-singleton domain. One cannot use the construction felicitously if only one individual is known in the context to satisfy the clausal predicate.

To account for separate-constituent licensing with either universal quantification or contrastive focus, as well as the use of both universal quantification and contrastive focus in the same clause, I treat each unpronounced licensor as a Focus head and allow for recursion to generate successively nested FocPs. Each head then checks a corresponding feature of either [+universal] or [+contrafocus] through an agreement operation, which is required because the pronouns mới and cũng would otherwise be interchangeable. An Agree operation (Chomsky 2001) may be applied in this case as the probe and the goal are in a c-command relation and occupy the same phase.

5. Particle Collisions

Let us look at the interaction of universal quantification and contrastive focus, both of which may appear in the same clause, in which case they produce scattering effects as the particles collide at the edge of TP. We can take a transitive verb with an agentive subject and apply contrastive focus to either the subject or the object and universal quantification to the other argument. Furthermore, universal quantification and contrastive focus have two possible relative scopes with differing interpretations. When we test all four of these combinations, we find that in each case, the particles cũng and mới appear side by side preverbally but not in the same order for all combinations.16

---

15 It may be possible to extend this account to Mandarin and dialects of Vietnamese that do not permit separate-constituent licensing if a parameter in these languages prohibits recursion on FocP.
16 Unfortunately, the consecutive use of cùng and mới in either order renders a sentence unacceptable for my consultants, so the argument from interacting particles cannot be made using independent data,
(18) Contrastive subject > universal object
   a. * [ Chỉ đứa này ] [ món nào nó thấy ] CUNG mooie ăn.
      only child this dish which 3SG see CUNGMOI eat
   b. [ Chỉ đứa này ] [ món nào nó thấy ] mooie CUNG mooie.
      only child this dish which 3SG see MOI CUNG Geat
   ‘It was only this child who ate every dish s/he saw.’

(19) Contrastive object > universal subject
   a. [ Chỉ đứa này ] [ thầy nào của nó ] CUNG mooie phât.
      only child this teacher which POSS 3SG CUNGMOI punish
   b. * [ Chỉ đứa này ] [ thầy nào của nó ] mooie của CUNG mooie.
      only child this teacher which POSS 3SG MOI CUNG punish
   ‘It was only this child who was punished by all of his/her teachers.’

(20) Universal subject > contrastive object
   a. * [ Đứa nào ] [ chỉ bạn nó ] CUNG mooie theo.
      child which only friend 3SG CUNGMOI follow
   b. [ Đứa nào ] [ chỉ bạn nó ] mooie của CUNG mooie.
      child which only friend 3SG MOI CUNG follow
   ‘For every child x, it was only x’s friends that x followed.’

(21) Universal object > contrastive subject
   a. * [ Đứa nào ] [ chỉ thầy nó ] CUNG mooie la.
      child which only teacher 3SG CUNGMOI scold
   b. [ Đứa nào ] [ chỉ thầy nó ] mooie của CUNG mooie.
      child which only teacher 3SG MOI CUNG scold
   ‘For every child x, it was only x’s teacher who scolded x.’

It is not immediately clear what determines the varying orders observed for this scattering phenomenon. In each case, the operator scope order corresponds to order of the preposed constituents: if contrastive focus takes scope over universal quantification, then the focused constituent precedes the quantified wh-indeterminate, and vice versa. For some reason, whenever universal quantification takes scope over contrastive focus, the resulting sentence with CUNG và móc in the right order is somewhat degraded but remains more acceptable than the alternative with the particles in the wrong order.

The relative order of CUNG và móc, however, is independent of the scope order. Instead, the generalization is that if the subject receives contrastive focus, the order is CUNG preceding móc, but otherwise, móc precedes CUNG. Extending this test to intransitive arguments and adjuncts, we find the distribution in (22), based on sentences like the selection in (23).

---

at least to my knowledge. There are a number of possible reasons for why the particles cannot occur adjacently for these speakers, and this fact alone does not necessarily mean that CUNG và móc are not pronouns in their varieties of Vietnamese. Notably, these speakers allow CUNG và móc to be stressed for emphasis, whereas for me, CUNG và móc must remain unstressed and pronounced together with the following verb or tense, aspect, or modal head, which suggests that they are clitics. (When stressed for emphasis, CUNG has only the additive reading for me.) This is perhaps not too surprising given the syntactic parallels to Clitic Left Dislocation.
<table>
<thead>
<tr>
<th>Condition</th>
<th>Cùng mới</th>
<th>Mọi cuộc</th>
</tr>
</thead>
<tbody>
<tr>
<td>transitive subject(CF) &gt; transitive object(∀)</td>
<td>✓</td>
<td>*</td>
</tr>
<tr>
<td>transitive object(CF) &gt; transitive subject(∀)</td>
<td>✓</td>
<td>*</td>
</tr>
<tr>
<td>transitive subject(∀) &gt; transitive object(CF)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>transitive object(∀) &gt; transitive subject(CF)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>transitive subject(CF) &gt; adjunct(∀)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>adjunct(CF) &gt; transitive subject(∀)</td>
<td>✓</td>
<td>*</td>
</tr>
<tr>
<td>transitive subject(∀) &gt; adjunct(CF)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>adjunct(∀) &gt; transitive subject(CF)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>transitive object(CF) &gt; adjunct(∀)</td>
<td>✓</td>
<td>*</td>
</tr>
<tr>
<td>adjunct(CF) &gt; transitive object(∀)</td>
<td>✓</td>
<td>*</td>
</tr>
<tr>
<td>transitive object(∀) &gt; adjunct(CF)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>adjunct(∀) &gt; transitive object(CF)</td>
<td>✓</td>
<td>*</td>
</tr>
<tr>
<td>unergative subject(CF) &gt; adjunct(∀)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>adjunct(CF) &gt; unergative subject(∀)</td>
<td>✓</td>
<td>*</td>
</tr>
<tr>
<td>unergative subject(∀) &gt; adjunct(CF)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>adjunct(∀) &gt; unergative subject(CF)</td>
<td>✓</td>
<td>*</td>
</tr>
<tr>
<td>unaccusative subject(CF) &gt; adjunct(∀)</td>
<td>✓</td>
<td>*</td>
</tr>
<tr>
<td>adjunct(CF) &gt; unaccusative subject(∀)</td>
<td>✓</td>
<td>*</td>
</tr>
<tr>
<td>unaccusative subject(∀) &gt; adjunct(CF)</td>
<td>✓</td>
<td>*</td>
</tr>
<tr>
<td>adjunct(∀) &gt; unaccusative subject(CF)</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

(22) a. Contrastive transitive subject > universal adjunct

[ Chỉ đứa này ] [ ngày nào ] mọi cứng nghe nhạc AKB48.
only child this day which MOI CUNG listen music AKB48

‘It was only this child who listened to the music of AKB48 every day.’

b. Contrastive transitive object > universal adjunct

[ Chỉ cuốn này ] [ ngày nào ] tôi cứng mới đọc.
only book this day which I CUNG MOI read

‘It was only this book that I read every day.’

c. Contrastive unergative subject > universal adjunct

[ Chỉ đứa này ] [ ngày nào ] cứng mới nhảy múa.
only child this day which MOI CUNG dance

‘It was only this child who danced every day.’

d. Contrastive unaccusative subject > universal adjunct

[ Chỉ đứa này ] [ ngày nào ] cứng mới té.
only child this day which CUNG MOI fall

‘It was only this child who fell down every day.’

This shows that it is not just a transitive subject but rather an external argument—an unergative subject or an agentive transitive subject—that when receiving contrastive focus triggers the mọi cứng order. (Non-agentive transitives pattern with unaccusatives.)

If cứng and mọi are focal heads in the left periphery, it is difficult to explain why their order does not correspond to the expected scope order but is conditioned by their associated argument positions within the verbal domain. Even if we allow for such conditioning, the surface order of the constituents and the particles must be derived by moving the focal heads into the correct order, an operation that would appear to violate the Head Movement...
Constraint (Travis 1984), followed by subsequent fronting of the preposed constituents while preserving their scope order.

6. The Mechanics of Scattering and Fusion

Alternatively, cũng and mới can be treated as pronouns generated within the vP that raise to the edge of TP and merge through Predicate Abstraction to establish a link between an assignment to a variable and the position the variable is evaluated in (Heim and Kratzer 1998); this mechanism can also account for relative clauses. The pronouns must move past T, as both particles obligatorily precede tense, aspect, and modal markers, as shown in (24).

(24) a. [ Con quái vật Godzilla] [ thành phố nào ] mới cũng sẽ phá hoại.  
    ‘It is the monster Godzilla that will destroy every city.’

b. [ Con quái vật Godzilla] [ thành phố nào ] mới cũng đã phá hoại.  
    ‘It is the monster Godzilla that has destroyed every city.’

c. [ Con quái vật Godzilla] [ thành phố nào ] mới cũng phải phá hoại.  
    ‘It is the monster Godzilla that must destroy every city.’

We can suppose that Spec of TP consists of multiple landing sites and can accommodate all moved instances of cũng and mới, each one performing Predicate Abstraction as it merges. For semantic composition, the pronouns should merge in reverse order of the relative scope of the focal heads they are associated with, so that the application of each focal head to its predicate complement can assign values to the right variable. Thus, if contrastive focus takes scope over universal quantification, cũng must merge at a position higher than mới.

This order, however, is clearly not reflected in the surface order of cũng and mới. We can suppose that some mechanism forces all instances of cùng to precede all instances of mới, prior to any further extraction out of TP, resulting in an order of cùng mới. The mechanism could be a morphological process that orders the elements in Spec of TP upon spell-out, or it could be an agreement operation in the PF derivation that checks a special feature present on cùng but not mới, moving each instance of cùng to a higher agreement head above TP, as illustrated in (25b).

(25) a. [ Món đó ]i [ đứa nào ]j cùng j mới i ăn —i  
    ‘It is that dish that every child ate.’

b. 

- 224 -
Although not well motivated independently, this account may explain the degraded acceptability of sentences in which universal quantification takes scope over contrastive focus, as in (26). These cases would require extraction of cũng from under an intervening mới when both are in Spec of TP. If such an extraction is structurally dispreferred due to a superiority constraint, then the weakening in acceptability is not unexpected.

(26) a. ? [ Đứa nào ]_j [ cà ri ]_i cung _j mới _j ăn _i
‘For every child, it is curry that s/he ate.’

b. 

To account for the mới cũng order when an external argument receives contrastive focus, we can apply a Case checking analysis in which an external argument is assigned agentive Case but must check it in Spec of Agr\_S, above TP (Chomsky 1993). If Agr\_S is above the Agr head that checks the special feature on cũng, then extracting an external argument (whether it is mới, cừng, or an ordinary subject) to this position will always produce the observed word order. When mới moves, it ends up in a position preceding all instances of cừng, resulting in an order of mới cừng whenever a clause contains both contrastive focus and universal quantification, as in (27). This movement would in principle also apply to cừng when it is an external argument, though the movement is not detectable as cừng is already fronted in Spec of Agr. When an ordinary subject moves to Spec of Agr\_S, it precedes all instances of cừng and mới that may occupy any specifier positions below it, resulting in the observed word order where cừng and mới surface between the subject and the verb or tense, aspect, or modal marker.

17 There is an EPP feature on T that requires the specifier to be filled (i); in the case of an unaccusative verb, the object can therefore raise to this position and must do so when nothing else is available (ii). This position, however, does not check Case, a property that is consistent with the possibility of also satisfying the EPP by putting an adjunct in the specifier position in the absence of an external argument (iii).

(i) * Xuất hiện một trận bảo khổng lồ ở Fukuoka hôm qua. appear one CL storm giant at Fukuoka yesterday
Intended: ‘A giant storm appeared in Fukuoka yesterday.’

(ii) Một trận bảo khổng lồ xuất hiện ở Fukuoka hôm qua. one CL storm giant appear at Fukuoka yesterday
‘A giant storm appeared in Fukuoka yesterday.’

(iii) Hôm qua xuất hiện một trận bảo khổng lồ ở Fukuoka. yesterday appear one CL storm giant at Fukuoka
‘Yesterday, a giant storm appeared in Fukuoka.’
The Proceedings of GLOW in Asia IX


child this dish which MOI CUNG eat

'it is this child who ate every dish.'

b. 

AgrSP
AgrS' 
AgrP 
Agr' 

This explains the scattering effect, but what about fusion? To account for the appearance of only one instance of Cùng or mới in separate-constituent licensing, we can suppose that morphophonological haplology deletes extra adjacent instances of the same particle (28).


dish which child which CUNG CUNG eat

'Every dish is such that every child eats it.'

b. 

AgrSP
AgrS' 
AgrP 
Agr' 

There is in fact independent evidence for haplology. If the movements place two instances of the same pronoun in positions separated by an instance of the other pronoun, then the conditions for haplology would not be met, and we expect all three instances to be pronounced. This is indeed the case for constructions that produce a sequence of mới Cùng
mới: this particular sequence is only possible when two licensees receive contrastive focus, one of them being an external argument, and a third licensee receives universal quantification, such as in (29). The sentence can be uttered in a context in which some books are under discussion, and all but one of them were submitted by their authors’ literary agents to every publisher they knew. The interlocutor claims that the exception is Book A, for which it is the book’s author, and not the author’s agent, who submitted the book. The speaker then utters (29a) to say that it is not Book A but Book B that meets that condition.

\[(29)\ a. \quad [\text{Cuốn này }]_i [\text{tác giả nó}]_j [\text{nhà xuất bản nào họ biết }]_k \text{mới}_j\]

\[\text{cùng}_k \text{mới}_i \rightarrow j \text{gửi} \rightarrow k \rightarrow i\]

\['It is this book for which it is its author who sent it to every publisher they knew.\]

b. Movement of *cùng* to Spec of AgrsP does not produce any change in surface word order, so it is impossible to generate a sequence of *cùng mới cùng* under this account, and such sequences appear to be unattested. Nevertheless, (29) provides evidence for a haplology account, at least for *mới*.

7. Conclusion

The wh-universal quantification and contrastive focus constructions in Vietnamese pose a puzzle when they appear multiple times within the same clause. When two instances of the same focal operator appear, there is fusion, the reduction of two markers to a single pronounced particle, and when both focal operators are represented, there is scattering, the conditioned variation in the order of the particles *cùng* and *mới*.

In this paper, I argued that these phenomena can be accounted for by distinguishing the particles from unpronounced focal heads in the left periphery. This allows for fusion to be analyzed as an application of haplology, and for scattering to be analyzed as the result of Case checking. The analysis allows separate-constituent licensing in Vietnamese to be aligned with its counterpart in Japanese, and also accounts for non-adjacent licensing in other Vietnamese dialects and in Mandarin if recursion on Focus projections is blocked.
The shared syntactic properties of the universal quantification and contrastive focus constructions leave a number of questions that remain to be answered. What underlying syntactic and semantic properties are there that result in these two constructions patterning alike, and where do they belong in a larger account of the left periphery? How do the Vietnamese constructions relate to similar constructions in other languages, and what syntactic configurations are possible for expressing universal quantification or contrastive focus? Future work along these lines would illuminate the various components of information structure and the nature of the interactions among them.

References


ON THE SYNTACTIC LICENSING OF LOCATIVE EXPRESSIONS IN JAPANESE*

Yoshiki Ogawa¹ and Fumikazu Niinuma²
Tohoku University¹ and Morioka University²

1. Introduction

Since Vendler (1967) classified verbs into four semantic types, that is, states, activities, accomplishments and achievements, two major streams of theorizing this classification have been proposed. In one stream, Hale and Keyser (1993), Levin and Rappaport Hovav (1995), Rappaport Hovav and Levin (1998), and Kageyama (1996) propose that different verb meanings should be represented on their lexical semantic representations, a ‘submodule’ of the language faculty, with the basic assumption that the linguistically relevant information is projected from the lexicon, by a set of linking rules. In the other stream, the linguists of the minimalist syntax persuasion, including Mateu and Rigau (2002), Zubizarreta and Oh (2007), and Ramchand (2008), have explored a view of the architecture of grammar whereby the lexicon is eliminated as a module with its own special primitives and modes of combination, and argued for a syntactic representation of these basic classifications of verb meaning. Among them, Ramchand (2008) proposes to replace the lexical conceptual structure (LCS), as in (1a,b), by a purely syntactic representation, as in (2a,b), where InitP is the causing/initiating projection, ProcP is the process projection, and ResP is the result projection:

(1) a. \([\{x \text{ DO-something}\}]\) (activities)
   b. \([\{x \text{ DO-something}\} \text{ CAUSE } \{\text{BECOME } [y \text{ BE AP/PP}]\}]\) (accomplishments)

(2) a. \([\text{InitP} \text{ Init } \{\text{ProcP Proc DP (Rheme)}\}]\) (activities) (Ramchand (2008: 46))
   b. \([\text{InitP} \text{ Init } \{\text{ProcP Proc \{ResP NP [Res AP/PP]\}]\}]\) (accomplishments) (ibid.: 39)

According to Ramchand (2008: 39), postulation of InitP, ProcP and ResP is a splitting up of what we normally think of as VP, in the same spirit as Rizzi’s (1997) Split CP or Pollock’s (1989) Split IP, and the syntactic characterization of each projection is as described in (3):

(3) a. InitP represents the outer causational projection and is responsible for introducing the external argument. (Ramchand (2008: 39))
   b. ProcP is the heart of the dynamic predicate, since it represents change through time, and it is present in every dynamic verb. (ibid.: 40)
   c. The ResP only exists when there is a result state explicitly expressed by the lexical predicate. (ibid.: 40)
   d. The head of ResP is realized by the overt verb which can take a simple locational PP as its complement, and a DP in its specifier functions as the subject of the predicative PP. (ibid.: 75)

Independently, Thompson (2006) claims that, within the syntactic domain which Hale and Keyser (1993) identified as ‘lexical syntax’, there is at least one syntactic projection of

* We thank the audience at the conference for their valuable comments. This work was supported in part by a Grant-in-Aid for Scientific Research (C) (No.24520526). All remaining errors are our own.
the aspectual functional category that is responsible for the determination of the situation aspect. Thompson (2006) claims further that the time-frame adverbial, which modify a telic event, and the durative adverbial, which modify an atelic event, are syntactically licensed in AspP and VP, respectively. One more important claim by Ritter and Rosen (1998, 2000) is that a DP which contributes to the telicity of the entire event moves to [Spec, Asp], so that it has a different morphological realization from the one which does not contribute to the telicity of the entire event (e.g. Accusative Case vs. Partitive Case in Finnish).

In this paper, combining Ramchand’s proposals with the Thompson’s ideas and Ritter and Rosen’s ideas, we will make the following three assumptions:

(4) a. Between InitP and ProcP is the functional projection of AspP, whose head can bear the semantic feature of [+telic] and the formal feature of [+bounded], the latter of which must attract either or both of the Theme DP and the Locative PP with an interpretable [+bounded] feature (cf. Chomsky (1995)).

b. Although the unmarked occurrence of the Res has a [+telic, +bounded] feature and contributes to the telicity of the entire event, it can have a [+telic, -bounded] feature when it implies result state persistence (see section 3.2).

c. When the Res with a [+telic, -bounded] feature is merged with a [+bounded] DP, the ResP assumes the [+telic, +bounded] feature, whence it moves to [Spec, Asp].

Exploiting these assumptions, we will claim that, in Japanese, a locative expression which contributes to the telicity of the entire event and one which does not do it show different morphological realizations with the postposition ni and de, respectively. More specifically, we claim that when a locative PP (or the ResP containing it) moves to [Spec, Asp], it is morphologically realized as DP-ni, while it is morphologically realized as DP-de when it remains in ResP or is originally merged higher than AspP. Also, when a locative PP is merged with the Res with a [-bounded] feature or a stative predicate, too, it is morphologically realized as DP-ni. In short, The gist of our claim is that the morphological form of a locative expression in Japanese is exclusively determined by its syntactic position and the formal feature with which it is associated. Relevant examples are illustrated below:

(5) a. Taroo-wa heya-ni/*-de terebi-ga hosii.
Taroo-Top room-in television-Nom want
‘Taroo wants (to have) a television in his room.’

b. Taroo-wa heya-de/*-ni ne-tei-ru.
Taroo-Top room-in sleep-Prog-Pres
‘Taroo is sleeping in his room.’

We will defend our hypotheses from various facts in Japanese, including (i) the correlation between the telicity of the event denoted by the VP and the morphological form of a locative expression, (ii) a morphological distinction between resultative and depictive secondary predicates, (iii) the relative word order between an object-oriented floating quantifier and a durative adverbial denoting the temporal interval over which the result state

1 As for the question why the merger of Res head with a [-bounded] feature and a DP specifier with a [+bounded] feature can form a ResP with a [+bounded] feature which moves to [Spec, Asp], we assume that something very similar can be seen in the wh-movement such as Whose father did you see there? In fact, in the movement of ResP to [Spec, Asp], the feature of the complement of Res is also relevant, because the resultative construction with a [-bounded] AP is interpreted as an atelic event:

(i) a. I laughed myself sick in three hours/*for three hours. (Tenny (1994: 43))
   b. Mary hammered the metal flatter and flatter for three hours. (Travis (2010: 110, note 27))
of a delimited event can persist, (iv) the fact that a resultative predicate occurs closer to the verb than a locative expression even if they both are headed by *ni* (Takezawa (2000)), and (v) the failure of a locative *ni*-phrase in nominal projections, among others.

This paper is organized as follows: in section 2, we will introduce basic facts about the distribution of locative *ni* and locative *de* in Japanese, certain differences between resultative and depictive secondary predicates, and the failure of the locative *ni* to occur in the environments of nominalization contexts. We will summarize how Nakau (1994, 1995) and Takezawa (2000) explain some of these facts and why they cannot explain the others. In section 3, we will introduce a couple of syntactic devices that we will adopt in this paper, including those introduced in (4), present some pieces of evidence for them, and provide a syntactic explanation of the basic facts introduced in section 2. Section 4 is a conclusion.

2. Japanese Data To Be Explained

2.1. Locative *Ni* and Locative *De*

In English, although there are various kinds of locative prepositions, those with the same meaning may be marked by the same preposition, whether they cooccur with a stative predicate or a dynamic predicate. Thus, *at, in, on,* and *to* can occur in both dynamic and stative predicates. By contrast, in Japanese, there is a well-known asymmetry between dynamic and static predicates in regard to the morphological realization of a locative expression. Shibatani (1977) makes the generalization that locative expressions which modify a dynamic predicate is marked by *de*, whereas those which modify a static predicate is marked by *ni*. His statement about static predicates seems exceptionless, as shown in (5a). However, the fact about dynamic predicates is not that simple. (5b) shows that the dynamic predicate *odoru* ‘dance’ can only accommodate a locative DP-*de*. However, first, as shown in (6a,b), some dynamic predicates allow only a locative postposition *ni* on one of its locative DP, while they allow only a locative postposition *de* on another locative DP. Second, as shown in (7a,b), some dynamic predicates allow only the locative postposition *ni* on its locative DP:

(6) a. Genkan-de/*-ni  atama-ni/*-de  boosi-o  kaburu.  
entrance-at    head-on    hat-Acc    wear
‘I wear the hat on my head at the entrance.’

b. Narita-kuukoo-de/*-ni  Taroo-ni/*-de  atta.  
Narita-airport-at     Taroo-to/at    met
‘I met Taroo at Narita airport.’

(7) a. Gakkoo-ni/*-de  toochaku-sita.  
school-at    arrived
‘I arrived at school.’

b. Gakkoo-ni/*-de  kasa-o  (oki-)wasureta.  
school-at    umbrella-Acc   (put-)forget
‘Inadvertently, I left my umbrella at school.’

Third, and most interesting for our concern, some dynamic predicates allow both a locative *ni* and *de* on the same argument DP, as shown in (8a-d):

(8) a. Sono fune-wa  hokkyoku-no  umi-ni/-de  kieta.  
that ship-Top arctic-Gen    sea-NI/-DE    disappeared
‘That ship disappeared in the Arctic Ocean’
Niwa-ni/-de kosumosu-ga saita
flower-NI/-DE cosmos-Nom bloomed
‘Flowers of cosmos bloomed in the garden.’

Taroo-wa niwa-ni/-de mizu-o maita.
Taroo-Top garden-NI/-DE water-Acc sprinkled
‘Taroo sprinkled water on the garden.’

Sono sakana-wa sinkai-ni/-de seisoku-suru.
that fish-Top abyss-NI/-DE inhabitation-do
‘That fish inhabits in the abyss.’

Nakau (1995: 21) proposes the invisible kyokumen doosi or ‘phasal verb’, which feeds the meaning of ‘A situation happens somewhere.’ In the spatial cognitive structure he proposes, the situation denoted by the phasal verb dominates another situation expressed by the basic predicate, as in (9), and the outer PLACE that is the argument of the phasal verb is manifested with de, while the inner PLACE that is one of the arguments of the basic predicate is followed by ni:\(^2\)

\[(9)\]
\[
\begin{array}{c}
\text{SITUATION2} \\
\text{SITUATION1} \\
\text{THING} \\
\text{PLACE}_{\text{OUTER}} (de, etc) \\
\text{PLACE}_{\text{INNER}} (ni, etc)
\end{array}
\]

Importantly, Nakau claims that, since the structure in (9) is the spatial cognitive structure which is a reflection of the human cognitive capacity, this distinction between the two types of locative expression is not unique to Japanese but can be extended to English as well. He claims that, although there is no morphological reflection in English, the distinction has syntactic reflections on the word order, as in (10a,b), and extraction possibility, as in (11b,c):

\[(10)\]
\[
a. \text{Many people drink in pubs in London.} \\
b.* \text{Many people drink in London in pubs.}
\]

\[(11)\]
\[
a. \text{I slept in my bed in New York.} \\
b. \text{Which bed did you sleep in in New York?} \\
c.* \text{Which city did you sleep in your bed in?}
\]

Notice that Nakau’s (1994, 1995) spatial cognitive structure can eventually be subsumed into Ramchand’s (2008) syntactic three-layered VP-structure. More specifically, since Nakau divides the basic predicates into action, state, process, where he identifies the process as spatial displacement relation, it would be more reasonable to make the semantic category of SITUATION1 and SITUATION2 in (9) correspond to the syntactic category of ProcP (rather than ResP) and InitP, respectively. Thus, Ramchand’s syntactic structure can neatly accommodate what Nakau’s spatial cognitive structure tries to capture, without adopting it.

Unfortunately, one potential problem with Nakau’s theory is that it cannot explain why the same process predicate can sometimes take a locative expression headed by ni and otherwise by de, as in (8a-d).

A closer examination of the ni/de alternation reveals that the locative expression contributes to the telicity of the event denoted by the verb. Note also that a locative expression to which a holistic interpretation is forced, such as niwa-zentai-ni ‘garden-whole- NI’, can

\(^2\) Where the place argument is a cover term for the location argument and the end point argument.
only be headed by \textit{ni}, as in (12d) and (13c) (cf. the English data in (27) and (28)):

\begin{enumerate}
\item[(12) a.] Niwa-ni/de jup-pun mizu-o maita.  
\hspace{1cm} \text{garden- NI/DE ten-minute water-Acc sprinkled}  
\hspace{1cm} ‘I sprinkled water in the garden for ten minutes.’
\item[(12) b.] Niwa-ni/??de jup-pun-de mizu-o maita.  
\hspace{1cm} \text{garden- NI/??DE ten-minute-in water-Acc sprinkled}  
\hspace{1cm} ‘I sprinkled water on the garden \{for ten minutes /?? in ten minutes\}.’
\item[(12) c.] Niwa-zentai-ni jup-pun-de / *jup-pun mizu-o maita.  
\hspace{1cm} \text{garden-entire-NI ten-minute-in /ten-minute water-Acc sprinkled}  
\hspace{1cm} ‘I sprinkled water on the entire garden \{in ten minutes /for ten minutes\}.’
\item[(12) d.] * Niwa-zentai-de jup-pun-de / jup-pun mizu-o maita.  
\hspace{1cm} \text{garden-entire-DE ten-minute-in /ten-minute water-Acc sprinkled}  
\hspace{1cm} ‘*I sprinkled water on the entire garden in ten minutes / for ten minutes.’
\end{enumerate}

\begin{enumerate}
\item[(13) a.] Niwa-ni/?de is-shuukan kosumosu-ga saita.  
\hspace{1cm} \text{garden- NI/?DE one-week cosmos-Nom bloomed}  
\hspace{1cm} ‘Flowers of cosmos bloomed in the garden for a week.’
\item[(13) b.] Niwa-ni/??de, uete-kara is-shuukan-de kosumosu-ga saita.  
\hspace{1cm} \text{garden-NI/??DE plant-after one-week-in cosmos-Nom bloomed}  
\hspace{1cm} ‘Flowers of the cosmos bloomed in the garden in a week after I plant them.’
\item[(13) c.] Niwa-zentai-ni/#de is-shuukan-de kosumosu-ga saita.  
\hspace{1cm} \text{garden-entire-NI/#DE one-week-in cosmos-Nom bloomed}  
\hspace{1cm} ‘Flowers of the cosmos I planted bloomed in the garden in a week.’
\end{enumerate}

A generalization at the first approximation seems to be as in (14):

\begin{enumerate}
\item[(14)] When a locative PP that modifies a dynamic predicate is morphologically realized as either the \textit{ni}-phrase or the \textit{de}-phrase, only the former can contribute to the delimitation (or “measuring out”; Tenny (1994)) of the event.
\end{enumerate}

Nakau’s (1994, 1995) theory cannot explain this generalization, since he has not argued anything about the correlation between SITUATION1 and telicity / delimitation of an event.

An important fact arguably related to the generalization in (14) is that lexically specified atelic verbs in Japanese cannot take a locative PP headed by \textit{ni} ‘at/to’ that refers to the end point of the motion, although they are compatible with a directional PP headed by \textit{ni-mukatte} ‘toward’, as in (15a,b). The difference between \textit{nimukatte} ‘toward’ and \textit{ni} in terms of telicity basically corresponds to the contrast between \textit{toward} and \textit{to} in English, as in (16a,b):

\begin{enumerate}
\item[(3)] The durative adverbial is possible in (12a), where the locative \textit{ni} is used, because the Accusative Case-marked object \textit{mizu-o} ‘water-Acc’ can have either a bounded or an unbounded interpretation. If we coerce a bounded interpretation on \textit{mizu} as in (i), the durative adverbial is excluded.
\item[(i)] Taroo-wa niwa-ni ip-pun-de / *ip-pun baketsu-jup-pai-no mizu-o maita.  
\hspace{1cm} \text{Taroo-Top garden-entire-on one-minute-in / one-minute bucket-ten-CL-Gen}  
\hspace{1cm} water-Acc sprinkled  
\hspace{1cm} ‘Taroo sprinkled ten buckets of water to the garden \{in a minute /for a minute\}.’
\end{enumerate}

\begin{enumerate}
\item[(4)] This sentence is acceptable with the durative adverbial, if it refers to the situation in which there is a large enough garden to the whole of which many sprinklers can spray water at the same time and the water-sprinkling continued for ten minutes. However, even in this situation, (12d) is unacceptable.
\end{enumerate}
(15) a. Taroo-Top station-to ran
   ‘Taroo ran to the station.’
   b. Taroo-Top station-toward three hour ran

(16) a. Mary ran towards the store for 3 hours/*in 3 hours.
   b. Mary ran to the store in 3 hours/*for 3 hours. (Travis (2010: 110))

If SITUATION1 could always license a locative ni-phrase, we would be unable to explain the ill-formedness of (15a). If, on the other hand, we assume that the locative expression that modifies a dynamic predicate is realized as the ni-phrase only when it is merged in ResP (and moves to [Spec, Asp]), then we can exclude (15a) because, under Ramchand’s (2008) proposal, a lexically atelic verb does not have the projection of ResP (see (3c)).

2.2. Resultative Ni and Locative Ni

Despite these apparent problems with Nakau’s theory, one aspect in which Nakau’s claim sounds intuitively correct is the following fact about resultative and depictive secondary predicates: (i) the former can be marked by ni but not by de, while the latter can be marked by de but not by ni, as shown in (17), and (ii) hierarchically, the former occurs in a lower position than the latter, as shown by the contrast in (18):

(17) a. Taroo-Top car-Acc deep red-NI/de painted
   ‘Taroo painted the car deep red.’
   b. Taroo-Top naked-DE/-NI danced
   ‘Taroo danced naked.’

(18) a. Taroo-Top car-Acc naked-DE deep red-NI painted
   ‘Taroo painted the car deep red naked.’
   b.* Taroo-Top car-Acc deep red-NI painted

In Nakau’s theory, the resultative secondary predicates could be licensed in the same position as PLACEINNER in (9), whereas the depictive ones could be licensed in the same position as PLACEOUTER in (9).

Takezawa (2000) argues, however, that the locative and resultative ni-phrases should be licensed in syntactically different positions, because the distribution of floating quantifiers (FQs) in terms of the two types of ni-phrases shows the following contrast:

(19) a. Taroo-Nom paint-Acc car-on three-kind painted
   ‘Taroo painted cars with three kinds of paints.’

5 As for (16b), we assume, following Mateu and Rigau (2002) and Zubizarreta and Oh (2007), that ran is adjoined to the phonetically empty telic verb GO in the syntactic template of [x GO to the store], an option unavailable in Japanese for some reason.
   parcel-Nom home-to three-CL arrived
   ‘Three parcels arrived home.’

(20) a.*? Taroo-ga penki-de kuruma-o makka-ni san-dai nutta.
   Taroo-Nom paint-with car-Acc really red-NI three-CL painted
   ‘Taroo painted on three cars really red.’

b.*? Syatsu-ga dorodarake-ni san-mai yogoreta.
   shirt-Nom muddy-NI three-CL got-dirty
   ‘Three shirts got muddy.’

Given Miyagawa’s (1989) mutual-c-command analysis of FQs in Japanese and the wide-spread assumption that Nominative subject of the unaccusative verbs is initially merged VP-internally, Takezawa argues that the locative ni-phrase is base-generated in a position higher than the underlying position of the Theme DP, which moves across it, whereas the resultative ni-phrase is base-generated in a position lower than the underlying position of the Theme DP, as in (21a,b):

(21) a. \[ VP \text{Locative Expression} [\text{V' DP (Theme) FQ V}] \]

b. \[ VP \text{DP (Theme) FQ} [\text{V' Resultative AP/PP V}] \]

Since the FQ associated with the Theme DP can never occur in a position lower than the resultative ni-phrase at any point in the derivation, (20a,b) will be ruled out.

Alternatively, if we are to update Takezawa’s proposal in (21b) in terms of Ramchand’s (2008) three-layered VP-structure, we can assume that the resultative predicate is directly merged with the head of ResP, the Theme DP is originally merged at the Spec of the ResP, and the ResP containing both is moved to [Spec, Asp], whereas the ni-marked locative PP which is similarly merged with the head of ResP can move alone to [Spec, Asp], as in (22a,b):

(22) a. \[ DP_i (\text{Theme}) … \{\text{AspP Locative PP} \}_{\text{ResP}} [\text{ProcP} t_i \text{FQ} \{\text{Res'} t_i \text{Res}\}] \]

b. \[ DP_i (\text{Theme}) … \{\text{AspP} [\text{ResP} t_i \text{FQ} \{\text{Res'} \text{Resultative PP Res}\}]_{\text{Asp'}} [\text{ProcP} t_j \text{Proc} \text{Asp}] \]

2.3. Nominalization

A third characteristic of a locative ni-phrase is that it is always excluded in the immediate projection of a nominal category: when ni can alternate with de in a verbal phase, it must be disambiguated to de in its nominal counterpart. The generalization is stated in (23), and for its typical illustrations, compare (8b) with (24) and (8d) with (25):

(23) A locative ni-phrase is always excluded in a nominal phase.6

(24) niwa-de-no/*niwa-ni-no kosumosu-no kaika
garden-DE-GEN/*garden-NI-GEN cosmos-GEN blooming
   ‘the blooming of a cosmos in the garden’

(25) sono sakana-no sinkai-de-no/*sinkai-ni-no seisoku
that fish-GEN abyss-DE-GEN/abyss-NI-GEN inhabitation

---

6 We are using the term “a nominal phase,” rather than “a nominal context,” because a locative PP in a nominal context can be headed by ni, if it occurs in a verbal phase (i.e. vP or CP) embedded in a noun phrase, such as a relative clause or a postsyntactic compound (Shibatani and Kageyama (1988)).
‘The inhabitation of the fish in the abyss’

2.4. A Summary

We have observed four generalizations about the locative ni- and de-phrases: (i) stative predicates are only compatible with the ni-phrases; (ii) when a dynamic predicate allows both ni-marked locatives and de-marked ones, they differ in terms of telicity; (iii) although both the locative and resultative ni-phrases contribute to the telicity, the former can c-command the underlying position of a theme NP at a point in the derivation, while the latter cannot; (iv) in a nominalized context, a ni-marked locative is always excluded. Note also that in English there is no morphological distinction between the locative expressions modifying a dynamic predicate and those modifying a static one. Given these facts, it is reasonable to claim that the distribution of a locative ni-phrase in Japanese is constrained morphosyntactically at least.

In the next section, we will claim that there are two syntactic positions for the licensing of a locative ni-phrase: one is within Ramchand’s (2008) ResP with a [-dynamic, -bounded] feature (in the case of a locative ni-phrase associated with a static predicate), and the other is the Spec of a functional category Aspect with a [+dynamic, +telic] feature. In the latter case, we assume that either the locative PP itself or the ResP containing it must be raised to [Spec, Asp] in order to check off an uninterpretable [+bounded] feature of Asp.

3. The Proposals and Assumptions

3.1. Morphosyntactic Manifestations of Telicity and/or Boundedness

We have assumed, in the spirit of Ramchand (2008), that an accomplishment verb occurs in the three-layered syntactic projection including InitP, ProcP, and ResP, whereas an activity verb occurs in the two-story syntactic structure including only InitP and ProcP. We have also assumed, following Thompson (2006) and Travis (2010), that there is an aspectual functional category between the underlying position of the external argument and that of the internal argument, which we will call “(Inner) AspectP” here. Combining these two assumptions, we will have the structures in (26a,b) for sentences denoting telic and atelic events, respectively:

(26)  a. accomplishments (telic events):
\[
[\text{InitP} \text{DP (Agent)} [\text{Init} \text{[AspP Asp (+telic)} \text{ProcP}[\text{ProcP}[\text{ResP} \text{DP (Theme)}]]]]
\]
\[
[\text{Res'} \text{Res (V(telic)} \text{AP (Stative/Resutative) / PP (Locative)))]]
\]

b. activities (atelic events):
\[
[\text{InitP} \text{DP (Agent)} [\text{Init} \text{[AspP Asp (-telic)} \text{ProcP}\text{[ProcP[ProcP}[\text{ProcP DP (Rheme)} (\text{PP (Locative))})]]
\]

With these structures presupposed, we will propose that the locative PP in (26a), but not in (26b), moves to [Spec, Asp] when it contributes to the telicity of the entire event, and that it is morphologically realized as DP-ni when it moves to [Spec, Asp], and as DP-de otherwise.7

It is not uncommon across languages that an argument DP, whether it is Location or Theme, differs morphosyntactically when it makes an entire event telic and when it does not. First, in English, Ritter and Rosen (1998: 141) point out that a locative DP in a sentence-final PP does not measure out the event so that it is morphologically unconstrained (cf. Tenny (1994)), whereas a locative DP immediately following the main verb must measure out the event so that it has to accompany an adjective such as whole/entire, which contributes to the holistic interpretation of the DP. Relevant examples are shown in (27) and (28):

7 As for the cases of result state persistence, as a special case of (26a), see section 3.2.
(27) a. The children taped pictures onto the wall.
b.* The children taped the wall with pictures.
c. The children taped (up) the whole wall with pictures.

(28) a. The tailor sewed buttons onto the dress.
b.* The tailor sewed the dress with buttons.
c. The tailor sewed (up) the entire dress with buttons.

We interpret the contrast between sentence (a) and sentences (b,c) in (27) and (28) as showing that the adjectives such as whole/entire are a morphological manifestation of the fact that a locative NP has been moved to [Spec, Asp] in order to check the [+bounded] feature of Asp.

Second, there are many cases in which a different morphological Case or adposition on a Location or Theme argument affects the telicity of the event denoted by the verb phrase. (16a,b), repeated below as (29a,b), are examples of English, taken from Travis (2010), and (30a,b) are examples of Finnish, taken from Ritter and Rosen (2001):

(29) a. Mary ran towards the store for 3 hours/*in 3 hours. 
b. Mary ran to the store in 3 hours/*for 3 hours.  (Travis (2010: 110))

(30) a. Anne rakensi talo-a tunni-n/*tunii-sssa.
Anne build house-Part hour-Acc/*hour-inessive
‘Anne was building a/the house for an hour/*in an hour.’ 
b. Anne rakensi talo-n vuode-ssa/*vuode-n.
Anne build house-Acc year-inessive /*year-Acc
‘Anne built a/the house in a year/*for a year.’  (Ritter and Rosen (2001: 436))

If the event delimiter moves to [Spec, Asp] and undergoes feature checking, it is not surprising that a DP or a PP which plays the same grammatical function shows a different morphological manifestation, depending on whether it measures out the event or not.

Third, the following Japanese examples show that, for some verbs, a locative DP-ni can measure out the event on its own, without a special aspectual marker on the verb, though a locative DP-o can measure out the event only if the verb is followed by a special aspectual marker, as the contrast between (31a) and (31b) shows (cf. Kishimoto (2001: 117)):8

Masao-Top wall -Loc poster- Acc tape-Past/tape-complete-Past
‘Masao covered posters onto the entire wall.’ 
b. Masao-wa kabe-o posutaa-de hari-tukushi-ta/*hat-ta.
Masao-Top wall-Acc poster-with tape-completive-Past/*tape-Past
‘Masao covered the entire wall with posters.’

We can interpret this asymmetry between (31a) and (31b) by assuming that, the locative ni-phrase is a morphological manifestation of the fact that it measures out the event, whereas an Accusative Case-marker on a locative DP is not. It follows from this that, in the former case, the telicity does not have to be doubly indicated by the overt aspectual marker tukusu ‘complete’, whereas the aspectual marker is obligatory in the latter case.

8 Kageyama (1993) identifies tukus ‘complete’ as a lexical verb. On the other hand, Nishiyama and Ogawa (2011) and Ogawa and Niinuma (2011) argue, in the spirit of Fukuda (2007), that tukus is an aspectual functional category that selects vP as its complement. Whichever approach will be taken for the exact nature of the “aspectual marker” does not affect our explanation of the facts in (31).
On the basis of these cross-linguistic facts, we claim that, just like the DP with whole/entire in (27c) and (28c), the PP LOC headed by to in (29b), and the DP Acc in (30b), the DP LOC-ni in (31a) is a morphological manifestation of the fact that it measures out the event so that it has been moved to [Spec, Asp]. Given this claim, de in (12b,d) is ruled out because the time-frame adverbial and zentai ‘whole’ are only compatible with a telic event, while the locative de-phrase associated with the verb maku ‘sprinkle’ is only compatible with an atelic event. Similarly, (12c) is ruled out with the durative adverbial (which denotes the temporal interval over which the action persists) because it is incompatible with the locative ni-phrase, which is a morphological manifestation of a telic event.

3.2. The Time-Frame Adverbial, the Durative Adverbial, and the FQs

All the accomplishment verbs and achievement verbs are compatible with a time-frame adverbial such as in an hour, due to their telic nature, as in (32a). A certain subset of them can also take a durative adverbial that specifies the temporal interval over which the result state of the event denoted by the telic verb persists, as in (32b). We will refer to (32b) as a sentence with the “result state persistence” reading of a telic verb (Kageyama (1996: 59)):

(32) a. In the magic show, the young lady disappeared in a second.
   b. The young lady disappeared for some time and then reappeared.

Given the VP structure in (26a), we will assume that a time-frame adverbial is adjoined to AspP, which lies between InitP and ProcP, as in (33a), and the durative adverbial is adjoined to either ProcP or ResP, as in (33b,c) (cf. Thompson (2006)). While adjunction of a durative adverbial to ProcP is only compatible with an atelic verb, its adjunction to ResP is only compatible with a telic verb, since an atelic verb should not have the projection of ResP. Moreover, we claim that adjunction of the durative adverbial to ResP, as in (33c), is possible when the event denoted by the telic verb implies the existence of a result state and the adverbial modifies the temporal interval over which it persists.

(33) a. InitP DP (Agent) [Init [[AspP Asp (+telic) [[ProcP Proc [[ResP DP (Theme) [Res’ Res (Vtelic) AP/PP]]] in an hour]]] (telic verbs with a time-frame adverbial)
   b. InitP DP (Agent) [Init [[AspP Asp (-telic) [[ProcP Proc (Vatelic) DP (Rheme) ] for an hour]]] (atelic verbs with a durative adverbial)
   c. InitP DP (Agent) [Init [[AspP Asp (+telic) [[ProcP Proc [[ResP DP (Theme) [Res’ Res (Vtelic) AP/PP]]] for an hour]]] (telic verbs with a durative adverbial)

A similar duality to the one in (32a,b) can also be observed in Japanese. Consider (34a,b), which show that the verb ireru ‘put’ taking a Theme DP and a locative ni-phrase is compatible not only with a time-frame adverbial but also with a durative adverbial which expresses the temporal interval over which the result state of the telic event persists:

(34) a. Kuruma-o san-pun-de shako-ni ireta.
car-Acc three-minute-in garage-in put
   ‘I put the car into the garage in three minutes.’
car-Acc one-hour garage-in put
   ‘I put the car in the garage for an hour.’ (Kageyama (1996: 57))

We will claim that (34a) and (34b) differ not only in the positions of the time-frame adverbial and durative adverbial but also in the position of the locative ni-phrase: the
ni-phase in (34a) is in [Spec, Asp], whereas that in (34b) remains in ResP. Evidence for this claim comes from an asymmetry in the distribution of floating quantifiers, as in (35) and (36):

car-Acc ten-CL three-minute-in garage-in put
'I put ten cars in the garage in three minutes.
b. Kuruma-o san-pun-de shako-ni jyuu-dai ireta.
car-Acc three-minute-in garage-in ten-CL put
c. Kuruma-o shako-ni san-pun-de jyuu-dai ireta.
car-Acc garage-in three-minute-in ten-CL put

car-Acc ten-CL one-hour garage-in put
'I put ten cars in the garage for an hour.
car-Acc one-hour garage-in ten-CL put
car-Acc garage-in one-hour ten-CL put

A generalization about the distribution of FQs with respect to the locative ni is stated in (37):

(37) a. When a time-frame adverbial modifies an accomplishment verb, an FQ that is associated with a Theme DP can follow a locative ni-phase.
b. When a durative adverbial specifies the temporal interval over which the results state of an event denoted by an accomplishment verb persists, an FQ that is associated with a Theme DP cannot follow a locative ni-phase.

Our explanation of this generalization is very simple. Suppose, in line with (4c) and (33c), that both the locative ni-phase and the durative adverbial in question are merged in the ResP. Moreover, suppose that the postpositional ni is a morphological realization of the fact that the locative PP has been moved to [Spec, Asp]. Then, (35c) will have the structure as in (38), overleaf. Here, the locative PP has moved across the FQ, which is stranded in the Spec of ResP when the Theme DP moves to a higher position (by scrambling or whatever syntactic operation). As a result, the surface word order of Theme DP-o > locative ni-phase > the time-frame adverbial > the object-oriented FQ follows.

On the other hand, in the structure of (36a) as in (39), overleaf, ResP is headed by a [+telic, +bounded] feature, which licenses the durative adverbial of the result state persistence reading, and the locative ni-phase here is licensed in situ. In fact, the locative ni-phase in (39) can NOT move to [Spec, Asp], since it must enter into a predication relation with the Theme DP in [Spec, Res], and its movement out of the ResP would break up the necessary predication. In this case, the uninterpretable [+bounded] feature on Asp can be checked against the interpretable [+bounded] feature of the entire ResP which can move to [Spec, Asp], as in (39) (cf. (4c) for the pied-piping of ResP). However, since the direct object (and the FQ) originates in [Spec, Res] and since both the locative PP and the durative adverbial are base-generated in a lower position in ResP, there would be no way to place the FQ following the locative ni-phase. This is why (36b,c) are ill-formed.

(feature percolation of <+bounded> from [Spec, Res] to ResP, ResP-movement to [Spec, Asp])

In short, the asymmetry between (35) and (36) follows from the following two assumptions: (i) whether the locative PP can move alone to [Spec, Asp] or the entire ResP must be pied-piped, and (ii) while the time-frame adverbial is licensed in AspP, the durative adverbial of result state persistence is licensed in ResP whose head has the [+telic, -bounded] feature.

3.3. Anaphor Binding

If the locative DP-\textit{ni} undergoes overt movement from inside ResP to [Spec, Asp] across the Theme DP in [Spec, Res], as we are claiming, then we predict that an anaphor contained in a locative DP-\textit{ni} can be bounded by the coindexed Theme DP even if the latter does not c-command the former in the overt syntax. On the other hand, we predict that an anaphor contained in an Instrumental DP-\textit{de}, which we can reasonably assume is adjoined to InitP, cannot be bounded by the coindexed Theme DP. This prediction is indeed borne out:

(40) a. Watashi-wa [otagai-no heya]-ni Taroo to Jiroo-o ire-ta.
   I-Top each other’s room-NI Taroo and Jiroo-Acc put-Past
   ‘I put Taroo and Jiroo into each other’s room.’

b.*Watashi-wa [otagai-no tsue]-de Taroo to Jiroo-o tatai-ta.
   I-Top each other’s stick-with Taroo and Jiroo-Acc hit-Past
   ‘I hit Taroo and Jiroo with each other’s sticks.’

Under our assumptions, the well-formedness of (40a) can be assimilated with that of (41a,b) (the so-called “connectivity effect”), and the ill-formedness of (40b) with that of (41c,d):

(41) a. John threw each other’s newspapers at Sue and Bill. (Pesetsky (1995: 222))

b. Each other’s remarks annoyed John and Mary. (ibid.: 43)

c.* I entrusted each other’s children with the adults in the room. (ibid.: 221)

d.* Each other’s teachers insulted John and Mary. (ibid.: 45)

3.4. Aspectuality and Nominalization

As the final piece of evidence for our proposals, let us return to the fact that locative \textit{ni} can never be licensed in a nominal phrase in Japanese. Relevant examples are (42a,b):

(42) a. Niwa-ni/de kosumosu-ga saita.
   garden-NI/DE cosmos-Nom bloomed (= (8b))

b. niwa-de-no/*niwa-ni-no kosumosu-no kaika
   garden-NI-Gen/*garden-NI-Gen cosmos-Gen blooming (= (24))

Given the proposals and assumptions we have made so far, we can attribute the ill-formedness of these sentences to the following assumption:

(43) In a nominal phase in Japanese, the AspP which could license a locative \textit{ni}-phrase is not permitted.
We will argue that this is not a universal syntactic prohibition but is a condition specific to the Japanese syntax. Note that event nominals in English are derived from a verb or an adjective by a nominalizing suffix. Grimshaw (1990) suggests that the complex event nominal, which she proposes as a subtype of process nominal which can be modified by an aspectual adjective, must realize its internal arguments obligatorily, as in (44):

(44)  a. The frequent expression *(of one’s feeling) is desirable.  (ibid.: 50)
    b. The constant assignment *(of unsolvable problems) is to be avoided.

Fu, Roeper and Borer (2001) provide a convincing argument that the process nominals, unlike the result nominals, are syntactically derived via the overt head-movement of a verb to a nominalizing suffix, as in (45a,b) (cf. also Ogawa (2001)):

(45)  a. \[ DP \[ NP N (-ation) \[ VP DP (Agent) \[ V^t V (investigate) DP (Theme)]]]]
    b. \[ DP DP (Agent) \[ D \[ NP V+N (investigation) \[ VP t_{DP (Agent)} \[ V^t t_v (of) DP (Theme)]]]]

Their evidence for this proposal is two-fold, as in (46a,b); evidence for (46a) and (46b) is illustrated in (47a) and (47b), respectively:

(46)  a. Only the process nominals can cooccur with VP-adverbs.
    b. Only the process nominals can be the antecedent of a do so anaphora, which must take a VP as its antecedent.

(47)  a. John's resignation suddenly and Bill's doing so too
    (resignation = process nominal derived from V)
    b.*John's trip and Bill's doing so too /*John's trip quickly
    (trip = underived nominal)

Combining Grimshaw’s (1990) observation and Fu, Roeper and Borer’s (2001) observations with our own proposals, we will have the full syntactic structure of a process nominal based on a telic verb, as in (48):

(48)  a. \[ DP \[ NP N (-ation) \[ InitP Init \[ AspP Asp \[ ProcP Proc \[ ResP NP \{ Res (examine) AP/PP\}]]]]
    b. \[ DP DP (Agent) \[ D \[ NP Res+Proc+Asp+Init +N (examination) \[ InitP t_{init} \[ AspP Asp \[ ProcP t_{proc} \[ ResP NP \{ t_{res} AP/PP\}]]]]

Here, we are assuming that Asp has an uninterpretable [+bounded] feature that needs to be checked against a [+bounded] feature of a DP or a PP (or the ResP containing both the DP and the PP), and that the PP is moved to [Spec, Asp] when it measures out the event denoted by the verb. In English, the checking is successful because a process nominal phrase in this language dominates AspP, to whose Spec the locative PP can move to check the

---

9 We assume that this pied-piping of ResP takes place in the case of resultative construction, too. Recall the contrast in (20), which shows that resultative AP/PPs do not move to [Spec, Asp] by itself. Note in this context that the resultative construction in Japanese cannot be nominalized either:

(i) a. \[ DP D \[ NP N (-ation) \[ V V (examine) DP (Theme)]]]
    b.*\[ DP DP (Agent) \[ D \[ NP Res+Proc+Asp+Init +N (examination) \[ InitP t_{init} \[ AspP Asp \[ ProcP t_{proc} \[ ResP NP \{ t_{res} AP/PP\}]]]]

- 241 -

- 241 -
uninterpretable [+bounded] feature of Asp (although there is no evidence that the checking takes place in English, since there is no morphological manifestation on the Locative PP).

On the other hand, a process nominal in Japanese is not morphologically deverbal but a verbal noun (VN), and hence it cannot be derived by concatenating a verbal projection with N. Note here that AspP is a verbal functional category. Then, it is natural to claim that, unless a VN is verbalized by the suffixal –suru ‘do,’ it cannot project up to AspP (or TP). This assumption enables us to explain the ill-formedness of (24) and (25) when the locative PP is headed by ni: the locative ni is excluded here because the Japanese DP headed by a VN does not dominate an AspP to whose Spec it could move; rather, (42b) has the structure in (49):

(49)   \[ DP \{VN, DP (kosumosu-no) \{VN PP (niwa-de-no/*niwa-ni-no) VN (kaika)\}\} D\]

By contrast, the locative de can be licensed in the same position in (49), because these phrases is not required to move to [Spec, Asp].

In short, we have argued that a ni-phrase is excluded in a nominal phase in Japanese, because the DP in Japanese does not dominate AspP, even if it is based on a process nominal. We can motivate this assumption from the fact that the process nominal in Japanese is not deverbal but is based on a verbal noun, which does not project up to Asp unless it is verbalized in advance.

4. Conclusion

Building on Ramchand’s (2008) first phase syntax, we have argued that in Japanese there are at least two different types of locative expressions which are licensed in different syntactic positions: one is the locative ni-phrase, which must be licensed in [Spec, Asp] or a stative (i.e. [-bounded]) ResP, and the other is the locative de-phrase, which can be licensed elsewhere. Moreover, we have suggested that the elements which are licensed in [Spec, Asp] are not limited to the locative ni-phrase but a resultative ni-phrase is another candidate, although the latter seems to pied-pipe the entire ResP containing it and the direct object DP in its Spec.

We have argued, as an extension of Thompson’s (2006) and Ramchand’s (2008) proposals, that while the time-frame adverbial is syntactically licensed in AspP, the durative adverbial is licensed in either ProcP or ResP, depending on whether it modifies the temporal interval over which the process continues or it modifies the temporal interval over which the result state created by a telic event persists. We have argued that the Res which licenses the durative adverbial or a resultative predicate is featurally specified as [+telic, -bounded], and the [+bounded] feature on the entire ResP is fed from the bounded DP in its specifier, so that the entire ResP moves to [Spec, Asp]. This dual nature of the featural specification of the head of ResP enabled us to explain an asymmetry in the (im)possibility of the word orders of Theme DP-o Locative DP-ni FQ V and of *Theme DP-o Resultative DP-ni FQ V.

Our proposals in this paper have three theoretical consequences. First, Thompson (2006) claims that a time-frame adverbial, which is only compatible with a telic event, is syntactically licensed in a position higher than where a durative adverbial is licensed. Similarly, our claim is that a locative ni-phrase, which measures out an event, is licensed in a position higher than where a locative de-phrase is licensed. These correlations between the temporal and locative adverbials could not be a coincidence, given Kratzer’s (1996) claim that the spatio-temporal arguments are licensed by one and the same event argument. Second, to the extent that the functional category of AspP plays an important role in Japanese too, we may have to reconsider Fukui and Sakai’s (2003) thesis that functional categories, if any, are syntactically inactive in Japanese. Third, there is the possibility to place our claim that the locative postposition ni is a morphological manifestation of telicity in the context of
“secondary grammaticalization” from Case markers to tense/aspect (TA) markers (cf. Hopper and Traugott (2003)). As for the tense marker, Evans (1995) suggests that some nominal Case markers in the Australian language Kayardild have a secondary use (called the “modal” use by Evans) in which their primary function is to indicate tense/mood. More specifically, temporal information in this language is sometimes encoded both by verbal inflection and by the modal proprietive/ablative Case marker on a DP, and otherwise (e.g. in the case of potential verbal inflection), exclusively by the modal proprietive/locative Case markers on a DP (and an adverb), as in (50):

(50)  a. Ngada kurri-\textit{nangku} mala-\textit{wu} (balmbi-\textit{wu}).  
     1sg.Nom see-Neg.POT sea-M.Prop tomorrow-M.Prop  
     ‘I won’t be able to see the sea (tomorrow).’

b. Ngada kurri-\textit{nangku} mala-\textit{y} (balmbi-\textit{y}).
     1sg.Nom see-Neg.POT sea-M.Loc tomorrow-M.Loc
     ‘I could not see the sea (yesterday).’

Nicolle (2012: 373) identifies these modal Case markers as a result of secondary grammaticalization from Case markers to tense/aspect (TA) markers. In a similar vein, we could argue that the locative postposition \textit{ni} that cooccurs with a telic predicate is identified with another instance of secondary grammaticalization from Case markers to TA markers. For the lack of diachronic evidence, however, we will leave an in-depth exploration of this possibility for future research.

References


1. Introduction

In this paper, I present the quantitative analysis of the patterns of repair strategies against marked segmental configurations in Japanese loanword phonology using a large-scale spontaneous speech corpus. I focus on the voiced obstruent geminates, and show how these marked configurations are resolved in spontaneous utterances. In the analysis, we have a closer look at the quantitative examination of the distribution of voiced geminates and repaired segments in terms of the chronological change and the phonological contexts. In doing so, I shed light on the properties of phonological factors that govern the distribution, such as Lyman’s Law, that play crucial roles in Japanese phonology.

This paper is organized as follows. In Section 2, I introduce the background information concerning the research topic based on some seminal works on voiced geminates, and clarify the problems and the goals of this research. Section 3 presents the method of analysis. In Section 4, I summarize the data collected. Section 5 presents the analysis. Finally, in Section 6 I conclude with the discussion.

2. Background

2.1. Research Topic

This research focuses on the voiced obstruent geminates (hereafter voiced geminates) that occur in Japanese loanwords, such as [bb], [dd], and [gg] (Kuroda 1965, among others). When loanwords with word-final voiced obstruents are borrowed into Japanese, the original voiced singletons are altered and realized as voiced geminates followed by an epenthetic vowel. As exemplified in (1), *dog* is realized as *doggu*, and similarly *bed* is realized as *beddo*.

(1) a. dog  =>  doggu   ‘dog’
    b. bed  =>  beddo   ‘bed’

However, voiced geminates are traditionally prohibited in Japanese phonology. Therefore, although voiced singletons are borrowed as voiced geminates, these segments should be avoided and undergo some phonological processes to fit into the well-formed segmental configurations in Japanese. The most common phonological process involved is the devoicing, as shown in (2) *doggu* undergoes devoicing and is realized as *dokku*, and similarly *beddo* is realized as *betto* with the voiced geminates being devoiced (Vance 1979; Itô and Mester 1986; Nishimura 2003; Kawahara 2006b et seq.).

(2) a. doggu =>  dokku   ‘dog’
    b. beddo =>  betto   ‘bed’

During the past half century, there has been a large influx of loanwords, and consequently the
items with voiced geminates are now widespread throughout the foreign stratum of Japanese lexicon (McCawley 1968; Vance 1987; Itô and Mester 1999), and thereby voiced geminates have diffused into Japanese phonology. Consequently, voiced geminates have become more likely to be regarded as a well-formed segment (Itô and Mester 1999). Still, voiced geminates are marked, and accordingly these segments are optionally repaired (change to less marked segments by phonological processes). This yields the variable phenomenon that consists of voiced geminates and repaired segments, as in [doggu]~[dokku], and [beddo]~[betto]. That is, the phonological processes as a repair strategy apply not categorically, but gradiently. Thus, the distribution of voiced geminates has to be characterized in terms of the probability.

2.2. Devoicing and Lyman’s Law

The devoicing of voiced geminates has long been subject to studies of various frameworks, such as Optimality Theory (Itô and Mester 1999; Nishimura 2003; Kawahara 2006, 2008), Harmonic Grammar (Pater 2009), and Phonetics – auditory perception (Kawahara 2011a, b; 2012b). Here, I review the findings of some seminal works on the devoicing of voiced geminates, and illustrate the relationship between the properties of devoicing and Lyman’s Law.

As mentioned above, devoicing is optionally applied as a repair strategy. The applicability of devoicing is partly predictable; in other words, there are patterns of devoicing. The most notable is Lyman’s Law (Itô and Mester 1986; Vance 2005), that has long been subject to the phonological research in Japanese, such as the research on rendaku (Vance 1979, 1980; Itô and Mester 1986, among others). In short, Lyman’s Law has the dissimilatory effects, and is regarded as a kind of OCP specific to the voice feature. That is, Lyman’s Law bans multiple occurrences of [+voice] within certain domains. Most of the examples of devoicing can be accounted for by Lyman’s Law. However, there are some patterned exceptions; in other words, some specific cases are subject to Lyman’s Law (Nishimura 2003).

(3) co-occurrence with other voiced obstruent(s) – devoicing

a. /doggu/ => [dokku]
b. /beddo/ => [betto]

(4) otherwise – no devoicing

a. /eggu/ => [eggu] (*[ekku]) ‘egg’
b. /reddo/ => [reddo] (*[retto]) ‘red’

(5) singleton – no devoicing

a. /gibu/ => [gibu] (*[gipu]) ‘give’
b. /bagu/ => [bagu] (*[baku]) ‘bug’

As shown in (3), if voiced geminates co-occur with other voiced obstruent(s), the devoicing applies as expected. For example, /doggu/ is realized as [dokku], and /beddo/ is realized as [betto]. However, if there is no other voiced obstruent, Lyman’s Law does not come into play, and devoicing is not triggered. Consequently, /eggu/ is not realized as *[ekku], and /reddo/ is not realized as *[retto] in (4). Furthermore, the voiced obstruent singleton never devoices, even if it co-occurs with other voiced obstruent(s), as shown in (5). Based on these observed

---

1 Deciding the domain of Lyman’s Law is one of the important tasks in the research on rendaku, as it is directly linked to the applicability of Lyman’s Law. The issue of domain is extensively discussed in Itô and Mester (2003).
patterns, Nishimura (2003) claims that neither voicing nor geminacy alone causes devoicing, but their combination can trigger the devoicing.

2.3. Problems and Goals

The previous studies on voiced geminates in Japanese loanwords have made substantial contributions to the phonological theory. Phonetic analysis has been conducted based on the empirical data, and the aspects of perception are being uncovered (Kawahara 2011a, b; 2012b). However, aspects of production are still understudied. In particular, the research using the spontaneous speech data has rarely been conducted. Furthermore, repair strategies other than devoicing have not been thoroughly studied. With this background in mind, the goals of this research are as follows. Primarily focusing on the production in spontaneous utterances, I shed light 1) on the gradient patterns manifested in the avoidance of voiced geminates, and 2) on the effects of phonological factors that govern the distributions of voiced geminates and repaired ones, such as Lyman’s Law.

3. Method

Data was retrieved from the Corpus of Spontaneous Japanese (CSJ, Kokuritsu Kokugo Kenkyuujo 2008). The CSJ is a large-scale spontaneous speech corpus of common Japanese with rich annotation (Sano and Hibiya 2012). In terms of its size, CSJ consists of 3,302 speech samples, amounting to 662 hours of speech, and 7.5 million words. CSJ mainly consists of two types of speech samples: One type is APS, representing Academic Presentation Speech that is stylistically careful/formal; the other type is SPS, representing Simulated Public Speaking that is stylistically relaxed/casual. CSJ provides rich annotations concerning speaker attributes and characteristics of each speech, such as style, nervousness, and spontaneity. This feature makes it possible to analyze external factors in detail.

Every relevant utterance of potential voiced geminates in the CSJ was extracted (i.e. all voiced geminates and repaired segments). An exhaustive search for the CSJ brought forth a total of 1,666 tokens of voiced geminates and repaired segments. Segmentation is based on the annotated transcription in CSJ. The transcription of the CSJ consists of two formats, one type is Kihonkei that is the written form, and the other type is Hatsuonkei that represents the phonetic form. If the utterance undergoes a phonological process, then Hatsuonkei represents the underlying form and the surface form separately; if the utterance is realized faithfully, then the underlying form and the surface form are represented by a single notation. By referring to this system, we can infer the phonological processes involved. The factors I consider in the analysis include: birth-year of speakers, type of repairs, type of voiced geminates, preceding/following vowels, and Lyman’s Law (specifically presence/absence of other voiced obstruent(s) within the same utterance). Many factors listed above are considered anew, because spontaneous utterances provide a detailed information about the actual usage, and enables to focus on the aspects that have been overlooked.

4. Data

This section presents the summary of the data. The extracted tokens were initially classified as one of two categories: either the segments are realized as voiced geminates, or the segments have undergone some sort of repair strategy. An example of the first category is an underlying /gg/ that is faithfully realized as [gg]. Two examples of the second category are an underlying /gg/ that undergoes devoicing and is realized as [kk], and an underlying /dd/ that undergoes degemination and is realized as [d]. This primary classification constitutes the basis of the following analysis. The probability of repair is calculated based on this variable.
Table 1 shows the overall distribution of voiced geminates and repaired segments.

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Probability (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>voiced geminates</td>
<td>1,153</td>
</tr>
<tr>
<td>repaired segments</td>
<td>513</td>
</tr>
<tr>
<td>total</td>
<td>1,666</td>
</tr>
</tbody>
</table>

As shown in Table 1, the frequency as well as the probability of voiced geminates are higher than those of repaired ones. Specifically, about 70% of all tokens are realized as voiced geminates that are faithful to the underlying representations. In other words, repair strategies are applied only 30% of the time. The results show that although it is traditionally claimed that voiced geminates are infrequent in Japanese phonology, many voiced geminates appear faithfully in production, especially in spontaneous utterances, in support of the claim that voiced geminates have diffused into Japanese phonology and have become more likely to be regarded as a well-formed segment (Itô and Mester 1999).

5. Analysis

With the overall distribution in mind, this section presents the analysis considering the factors that can affect the distribution of voiced geminates and repaired segments.

5.1. Chronological Change

I firstly demonstrate the chronological change of voiced geminates and repaired segments. As mentioned above, voiced geminates have penetrated into the Japanese phonology due to the influx of loanwords. Consequently, it may be the case that voiced geminates are becoming less marked over time, resulting in the gradual increase in the production of these configurations. In this section, I examine this possibility. The CSJ provides the information about the birth-year of speakers. Using this information, I conducted an apparent-time analysis, in which the difference in speakers’ birth-year corresponds to the apparent flow of time (see Bailey 2002 for apparent-time). I classified each token into six birth-year periods (grouped every 10 years): 1925-1934, 1935-1944, 1955-1964, 1965-1974, and 1975-1985. Based on this classification, I calculated the chronological transition of the probabilities of voiced geminates and repaired segments. Figure 1 illustrates the chronological changes in the probabilities of voiced geminates being repaired, according to the birth-year of speakers.

![Figure 1. The chronological change in the probability of repair](image_url)
As Figure 1 shows, the probability of repair rises as speakers’ birth-year becomes more recent.\(^2\) Itô and Mester (1999) claim that among the items in the Japanese lexicon, assimilated foreign items undergo devoicing (i.e., one of the repair strategies); on the other hand, unassimilated alien items remain intact without undergoing any repair. To interpret it broadly, the more the items are nativized into Japanese phonology, the more likely they are repaired. The increasing probability of repair in the present research suggests that the nativization of loanwords is developing over the years. Furthermore, given that the change proceeds not toward the acceptance of voiced geminates in Japanese phonology, but toward the alteration to fit into the Japanese template, we can argue that the voiced geminates still have the status of marked segments and are avoided.

### 5.2. Type of Repair

Next, I consider the distribution of repair strategies. As mentioned above, the most common phonological process involved in the avoidance of voiced geminates is the devoicing in the traditional characterization. However, repair strategies other than devoicing are also expected to apply. Here, we see what kinds of repair strategies are applied and the probabilities with which each repair is applied. Figure 2 shows the distribution of repair strategies.

<table>
<thead>
<tr>
<th>Repair Strategy</th>
<th>Frequency</th>
<th>Probability (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>devoicing</td>
<td>472</td>
<td>92.00</td>
</tr>
<tr>
<td>degemination</td>
<td>34</td>
<td>6.62</td>
</tr>
<tr>
<td>deletion</td>
<td>6</td>
<td>1.17</td>
</tr>
<tr>
<td>vowel lengthening</td>
<td>5</td>
<td>0.97</td>
</tr>
<tr>
<td>nasalization</td>
<td>3</td>
<td>0.58</td>
</tr>
<tr>
<td>rhotacization</td>
<td>1</td>
<td>0.19</td>
</tr>
<tr>
<td>total</td>
<td>521</td>
<td>100</td>
</tr>
</tbody>
</table>

As shown in Table 2, six types of repair strategies were observed. However, devoicing is predominant as it accounts for more than 90% of all repairs. The result supports the validity of the previous studies in the sense that the phonological process applied to the voiced geminates is predominantly devoicing. Furthermore, in most cases only a single repair strategy is applied, and only eight tokens show multiple repairs, such as the combination of devoicing and degemination.

I consider the uneven distribution of repair strategies in terms of economy. First, the predominance of devoicing may be associated with the least featural change (P-map hypothesis, Steriade 2001/2008). As exemplified in (6a), the featural change involved is limited to voice feature in devoicing; namely, [+voice] changes to [-voice] with place and manner features remaining intact. On the other hand, other repair strategies require a more broad change.

---

\(^2\) Loanwords in Japanese show the tendency to increase. The frequency of loanwords in the present data (loanwords with underlying voiced geminates) is gradually increasing, 1925-1944: 165; 1945-1964: 649; 1965-1984: 850, although the amount of utterance differs according to the age group, and a comparison based on the frequency may not yield an absolute measure.

\(^3\) In some cases, more than one repair strategy is applied in a single token. Therefore, the total frequency of repair strategies exceeds the number of tokens (513) where some kind of repair strategy is applied.
(6) a. devoicing: e.g. /beddo/ => [betto] (S00F0093)\textsuperscript{4} \\
change in voice feature dd => tt \\
[+voice] [-voice] \\
b. degemination: e.g. /bagudaddo/ => [bagudaado] ‘bagdad’ (S09M0855) \\
change in timing slot addo => ado \\
xxx => xx \\
compensatory lengthening ado => aado \\
xx => xxx \\

Taking degemination in (6b) as an example among other strategies, degemination itself renders the change in the timing slot, as in /dd/ => [d]. This leads to the additional change such as the compensatory lengthening, as in /bagudaddo/ changes to [bagudaado] with the preceding vowel being lengthened. Deletion, vowel lengthening, nasalization and rhotacization are exemplified in (7).

(7) a. deletion: loss of whole segment(s) \\
e.g. handoreddo => handore ‘hundred’ (S04F0792) \\
b. vowel lengthening: copy of whole segment(s) \\
e.g. gaabeddʒi => gaabeedʒi ‘garbage’ (A01M0103) similar to (6b) \\
c. nasalization: change in multiple features \\
e.g. neebaahuddo => neebaahundo ‘neighbourhood’ (A01M0565) \\
d. rhotacization: change in multiple features \\
e.g. disutoribyuuteddo => disutoribyuutoru\textsuperscript{5} ‘distributed’ (A03M0018) \\

d => r ([ɾ]) \\
[-sonorant] => [+sonorant] \\
[-nasal] => [+nasal] \\
[-delayed release] => \phi \\
[-continuant] => [+continuant] \\
[-delayed release] => \phi \\
[-approximant] => [+approximant] \\
[-tap] => [+tap] \\

Thus, the uneven distribution of repair strategies, i.e., the predominance of devoicing can be explained by the economy of the repair processes.

Next, the low frequency of multiple repairs is also associated with the economy. As mentioned above, voicing or geminacy alone is not marked, but the combination of these two creates the marked segmental configurations and can trigger the repair. By applying either one of voicing or geminacy, the marked configurations can be resolved. Therefore, the application of additional process(es) violates the economy and tends to be avoided, as it does not further contribute to the segmental well-formedness (cf. Kawahara 2011a for perception).

\textsuperscript{4} The alphanumeric character annotated to the end of each example (e.g. A04M0229) is the ‘speech ID’ that is used as the index of each speech. See Sano and Hibiya (2012) for more detail.

\textsuperscript{5} In this example, the alternation of vowels and degemination are involved; however, I focus only on the change from d to r.
5.3. Type of Voiced Geminate

In this section, I examine the effects of the type of voiced geminates. In the present data, five types of (underlying) voiced geminates are observed. These segments are characterized according to the place and manner of articulation. With respect to the manner of articulation, three types are observed: stops (/bb/, /dd/, /gg/), fricatives (/zz/), and affricates (/ddʒ/). According to the place of articulation, stops are classified into three types: bilabial, coronal, and velar. Fricatives and affricates are exclusively coronal. I hypothesize that the differences in segmental properties affect the probability of repair. The probability of repair by segment type is shown in Figure 2.

\[
\chi^2(4) = 94.45, p < 0.001
\]

Figure 2. The probability of repair by segment type

As Figure 2 shows, velar stops /gg/ are repaired with the highest probability; other segments are repaired in the following descending order: coronal stop /dd/, coronal fricative /zz/, coronal affricate /ddʒ/, and bilabial stop /bb/. Thus, in terms of the place of articulation the probability of repair shows a specific pattern. The pattern is summarized as the order in (8).

\[(8) \quad \text{/gg/} > \text{/dd/} > \text{/zz/} > \text{/ddʒ/} > \text{/bb/} \]

Velar sounds show the highest probability of repair whereas bilabial sounds show the lowest probability of repair. In comparison, coronal sounds show a probability that is intermediate between bilabial sounds and velar sounds. This suggests that the more back the place of articulation is, the more likely that the voiced geminates are repaired.

Next, I consider the geminate fricatives. Generally, geminate fricatives are perceptually non-prominent; and accordingly, singleton fricatives and geminate fricatives do not have the
sufficient distance in their contrast, compared with other sounds like geminate stops. Therefore, geminate fricatives are marked and avoided (Blevins 2004; Kawahara 2006a). In the present data, when underlying geminate fricatives undergo some phonological processes, all of these sounds undergo the affrication in addition to the devoicing.

(9) devoicing and affrication: e.g. /guzzu/ => [guttsu] (*[gussu]) ‘goods’ (S07M0392)

\[
devoicing \quad zz \quad \Rightarrow \quad ss \\
affrication \quad ss \quad \Rightarrow \quad tts
\]

As exemplified in (9), the voiced geminate fricative is realized as voiceless geminate affricate, for example, /guzzu/ is realized as [guttsu], and not as *[gussu]. The devoicing resolves the marked combination of voicing and geminacy, and yields the voiceless geminate fricative. However, the combination of geminacy and frication, that is also marked, still remains (see Ohala 1983; Lindblom and Maddieson 1988). Therefore, the additional process is called for, and in this case the affrication changes the fricative to the affricate. The second process appears to violate the economy principle if we only focus on the repair of the marked combination of voicing and geminacy; however, in order to avoid the marked geminate fricatives, the additional process is required.

5.4. Preceding Vowels

In this section, I examine the effects of vowels. First, I examine how the quality of preceding vowels affects the distribution of voiced geminates and repaired segments. In the present data, the voiced geminates and the repaired segments are observed in the environment of all five vowels of Japanese: /i/, /e/, /a/, /o/, and /u/.\textsuperscript{10} I plot the probability of repair for each vowel in Figure 3.

\[\chi^2(4) = 80.82, \ p<0.001\]

Figure 3. The effect of preceding vowels on the applicability of repair

As Figure 3 shows, the low vowel /a/ shows the highest probability; the high vowels /u/ and /i/ show the lowest probability; the mid vowels /e/ and /o/ show the intermediate probability between high vowels and low vowels. The distribution in Figure 3 can be summarized as the order in (10).

\textsuperscript{9} The order of phonological processes is out of the scope of this research.
\textsuperscript{10} In Japanese, the vowel /u/ is strictly /ɯ/ that is unrounded. However, I use the symbol u, as this difference does not relate to the purpose of the analysis.
As shown in (10), we can argue that the lower the preceding vowel, the higher the probability of repair. However, it is difficult to give a plausible phonetic account to this tendency, because in terms of aerodynamics the order of the probability of repair is the exact opposite of the difficulty in producing voiced geminates. The constriction of the vocal tract during the production of high vowels creates the greater air pressure in the oral cavity compared to other vowels, and this greater air pressure makes it difficult to produce voiced geminates (Ohala 1983; Hayes et al. 2004). Therefore, voiced geminates in the environment of high vowels should show the highest probability of repair. The reason why the lower preceding vowels induce more devoicing contrary to the aerodynamic consideration is an open question, and the examination of the effects of other factors that can contribute to the distribution here is also required. In any case, the result suggests that in the preceding context the application of repair is sensitive to height (or F1).

5.5. Following Vowels

Next, I examine how the quality of following vowel affects the distribution of voiced geminates and repaired segments. Among five vowels, only one token of voiced geminate with /e/ vowel that immediately follows the voiced geminate was observed, and no repaired counterpart was observed. Therefore, there can be a lexical gap here, and I removed the distribution in /e/ vowel from consideration. I plot the probability of repair for each vowel in Figure 4.

\[ \chi^2(3) = 63.95, p<0.001 \]

Figure 4. The effect of following vowels on the applicability of repair

As for the remaining vowels in Figure 4, the probability of repair rises from front vowel /i/ to back vowel /u/ as in (11).

(11) /u/ > /o/ > /a/ > /i/  
back  front

The result suggests that the more back the following vowels, the higher the probability. Therefore, we can argue that in the following context the application of repair is sensitive to
backness (or F2).\textsuperscript{11,12}

5.6. Lyman’s Law

In the following sections, I consider the effects of Lyman’s Law. Specifically, I examine the triggering factors of Lyman’s Law. As mentioned above, Lyman’s Law has the dissimilatory effects with respect to the voice feature, and bans multiple occurrences of \([+\text{voice}]\) within certain domains. Therefore, if Lyman’s Law is active in production, the probability of repair would be higher in contexts with violations of Lyman’s Law (see Kawahara 2012b and references cited therein for perception). The items I consider are summarized in (12).

\begin{enumerate}
\item Presence/absence of other voiced obstruent(s)
\item Locality: distance between voiced geminate and other voiced obstruent(s)
\item Number: the number of other voiced obstruents
\item Symmetry/asymmetry in preceding context/following context
\end{enumerate}

Firstly in (12a), I confirm the overall effects of Lyman’s Law by examining whether or not the presence or absence of voiced obstruents other than voiced geminates affects the probability of repair. Secondly in (12b), I examine whether or not the locality plays a role in determining the probability of repair. Specifically, I consider the distance between voiced geminate and other voiced obstruent(s). The distance is measured by intervening moras.\textsuperscript{13}

Thirdly in (12c), I examine whether or not the number of voiced obstruents other than voiced geminates plays a role in determining the probability of repair. Finally in (12d), I examine the symmetry or asymmetry in the preceding contexts and the following contexts, i.e., whether the Lyman’s Law has the same/different effects on the preceding contexts and the following contexts. In the analysis, I compare the preceding contexts and the following contexts in terms of the effects of Lyman’s Law. The analysis proceeds from the preceding contexts to the following contexts.\textsuperscript{14}

\textsuperscript{11} There is also an alternative account; namely, most of the \(/i/\) vowels are associated with affricates, producing \textipa{d\textaelji}, for instance, \textipa{jad\textaelji} ‘judge.’ Therefore, the distribution here can be biased, i.e., the distribution may be controlled by the place and manner features we examined in Section 5.3 other than vowels. If we ignore the distributions of \(/i/\) and \(/e/\) (recall that only one token of voiced geminate was observed preceding the vowel \(/e/\) and no repair was observed.), then the remaining distributions is limited to the \(/a/\), \(/o/\), and \(/u/\) vowels. The distributions that consists of these three vowels follows the pattern that is consistent with the aerodynamic consideration; namely, high vowels create the greater air pressure in the oral cavity, and the greater air pressure makes it difficult to produce voiced geminate, therefore, the probability of repair is highest in high vowels, and lowest in low vowels. The application of repair is sensitive to height (or F1); and contrary to the preceding vowels, we can give a straightforward phonetic account to this pattern.

\textsuperscript{12} Japanese has the open-syllable pattern (CV), and accordingly the voiced geminates are followed by vowels. Therefore, the voiced geminates and the following vowels are in the same syllable; on the other hand, the voiced geminates and the preceding vowels are in distinct syllables. This difference may be reflected in the direction and the degree of the effects of preceding/following vowels.

\textsuperscript{13} I examined the preceding and the following six moras from voiced geminates, because most of the Lyman’s Law-induced repair was caused by a trigger in adjacent moras (313/353=88.7%), and if the distance between voiced geminates and other voiced obstruent(s) exceeds two moras, the frequencies of both voiced geminates and repaired segments are significantly reduced. And in the distance of five moras, no repaired segment was observed.

\textsuperscript{14} The examination of the Lyman’s Law effects in the following contexts is presented in Section 5.6.4.
5.6.1. Presence/Absence of [+voice]

Lyman’s Law was shown to be active in perception (Kawahara 2012b). In this section, I examine whether or not the Lyman’s Law is also active in production. The effects can be tested by comparing the distributions with or without other voiced obstruent(s) than the voiced geminates. If Lyman’s Law is active, the probability of repair would be higher in “presence” than in “absence” of other voiced obstruent(s). In the analysis, any instance of voiced obstruent is regarded as “presence,” regardless of the locality and the number of voiced obstruents, since the purpose of the analysis here is to examine the overall effects of Lyman’s Law. The probability of repair according to the presence/absence is shown in Figure 5.

\[
\chi^2(1) = 330.76, \ p<0.001
\]

Figure 5. The probability of repair by presence/absence of [+voice] (preceding context)

As shown in Figure 5, the probability of repair is significantly higher in presence ([+voice], 48.8%) than in absence ([voice], 8.4%). The realization of voiced geminates tends to be blocked if there is other voiced obstruent(s). In other words, the presence of other voiced obstruent can induce the repair. Therefore, we can argue that Lyman’s Law is active also in production.

5.6.2. Locality

In the previous studies, Lyman’s Law was shown to be insensitive to the locality in perception (Kawahara 2012b). In this section, I examine the sensitivity of Lyman’s Law to the locality in production. If Lyman’s Law is sensitive to the locality, the probability of repair would be higher with voiced obstruents being closer to voiced geminates. In the analysis, I restrict our attention to the “presence” of other voiced obstruent(s), because the purpose of the examination is to see the effects of locality on the application of Lyman’s Law, and the Lyman’s Law does not play a role in the “absence.” Furthermore, I removed all tokens that contained more than one voiced obstruent, as in such cases, which obstruent triggers the devoicing is unclear. For example, in radifogiddo ‘Ladefoged’ (A01M0936) with two voiced obstruents: one is adjacent to the voiced geminate, and the other is three moras away from the voiced geminate, the distance between the voiced obstruents and the voiced geminate cannot
be uniformly determined.\(^\text{15}\) The distance is based on the number of intervening moras, including the voiced obstruent as shown in (13).

(13) \text{nee\textit{baahuddo}}
\begin{tabular}{ll}
\text{voiced obstruent} & \text{voiced geminate} \\
\end{tabular}
\begin{tabular}{l}
\text{intervening 2 moras} \\
\end{tabular}

The probability of repair by locality is shown in Figure 6.

![Figure 6. The probability of repair by locality (preceding context)](image)

As Figure 6 shows, the probability of repair is highest with the adjacent voiced obstruent, and it gradually declines as voiced obstruents move away from the voiced geminates. The result suggests that Lyman’s Law is sensitive to the locality in production, and the closer the distance is, the more likely the repair strategy is applied. Lyman’s Law is shown to have a stronger effect if the trigger and the target are closer together.

5.6.3. Number of voiced obstruents

In this section, I examine whether the Lyman’s Law is sensitive to the number of voiced obstruents in production. If Lyman’s Law is sensitive to the number, the probability of repair would be higher with more voiced obstruents. For the same reason in the previous discussion, I focused only on tokens that involve other voiced obstruent(s) than voiced geminates, where the Lyman’s Law is active. Example tokens for each number of voiced obstruents are given in (14).

(14) a. 1 voiced obstruent \text{sarahurredo} \text{‘thoroughbred’} (S00F0131)
b. 2 voiced obstruents \text{baundeddonesu} \text{‘boundedness’} (A02M0678)
c. 3 voiced obstruents \text{bagudaddo} \text{‘bagdad’} (S09M0855)

The distribution is summarized in Figure 7.

---

\(^{15}\) The multiple occurrence of other voiced obstruents in voiced geminates accounts for 10% (52/453) and the one in repaired segments accounts for 15% (82/449) of the entire “presence.” Therefore, the removal of the multiple occurrence does not significantly affect the result of the analysis.
Patterns in the Avoidance of Marked Segmental Configurations (S. Sano)

χ²(2) = 13.67, p<0.001

As Figure 7 shows, the probability of repair gradually rises from one voiced obstruent to three voiced obstruents. In other words, the more the number of voiced obstruents, the higher the probability of repair. The result suggests that the Lyman’s Law is also sensitive to the number in production, and its effect is incremental. The Lyman’s Law is shown to have a stronger effect if there are more triggers.

5.6.4. Symmetry/Asymmetry

Finally, I examine the symmetry/asymmetry of the preceding and the following contexts. As we have seen, in the preceding contexts Lyman’s Law was shown to be sensitive to all three items: the presence/absence, the locality, and the number of voiced obstruents. Here, I consider the relationship between the Lyman’s Law and these three items in the following contexts. If Lyman’s Law has the same effects on the preceding and the following contexts, then the distribution would be symmetric and there would be no difference between these two contexts; on the other hand, if Lyman’s Law has different effects on the preceding and the following contexts, the distribution would be asymmetric and there would be differences according to the contexts. Figure 8 through Figure 10 I show the distributions for each item in the following contexts.

---

16 Within six moras from voiced geminates, no token was observed where more than four voiced obstruents were involved.
As Figure 8 shows, the probability of repair does not differ according to the presence or absence of other voiced obstruent(s). Similarly in terms of the locality in Figure 9, the
distribution does not show a specific pattern. Only the number of voiced obstruents (Figure 10) shows a statistically significant pattern. The probability of repair is affected by the number of other voiced obstruent(s). Therefore, similar to the preceding contexts, the Lyman’s Law is sensitive to the number of voiced obstruents in the following context as well as in the preceding context.

However, the direction of the effect is reversed; namely, the more voiced obstruents, the lower the probability of repair. The result suggests that in the following context the number of voiced obstruents has an anti-Lyman’s Law effect. In summary, the preceding and the following contexts are under the asymmetric effects of Lyman’s Law.

6. Conclusion

In this paper, I focused on the voiced geminates in Japanese loanword phonology and the manner by which these marked segmental configurations are avoided. I examined the patterns of avoidance focusing on the aspects of production, and the phonological factors that contribute to the patterns. The findings of this research are summarized as follows.

Firstly, repair strategies are becoming more likely to apply over the years. This suggests that the loanwords containing voiced obstruents are gradually being nativized. Secondly, among the repair strategies, devoicing is predominant. The choice of the strategies is dependent on the economy, such as the “least featural change” and the “minimal repair.” Thirdly, in terms of the type of voiced geminates, the more back the place of articulation is, the higher the probability of repair is. The combination of the voicing and the geminacy is marked and avoided by the devoicing. The geminate fricative is also marked and avoided by the affrication. Finally, repair strategies are sensitive to the height (or F1) of the preceding vowel, but to the backness (or F2) of the following vowel. To take the results of the place of articulation of voiced geminates and in the following vowels together, we can argue that the likelihood of repair is dependent on the backness.

Lyman’s Law is sensitive to the presence/absence of other voiced obstruent(s), the locality, and the number in the preceding contexts. Firstly, if other voiced obstruent exists, then voiced geminates are more likely to be repaired. Secondly, as for the locality, the closer the voiced obstruents, the higher the probability of repair. Thirdly, as for the number of voiced obstruents in the immediate vicinity, the more voiced obstruents there are, the higher the probability of repair is. In the following contexts, only the number of voiced obstruents plays a role in the following manner: the more voiced obstruents there are, the lower the probability of repair is. Thus, the effect of the preceding context and the following context is asymmetrical. Specifically, Lyman’s Law is more sensitive to the preceding contexts. This suggests that in the flow of speech Lyman’s Law looks ahead and has progressive effects.

I conclude by mentioning an issue to be addressed in the future. The explanation of some findings in this research is left open. The theoretical and/or empirical examination of the possible effects of factors as well as the interaction thereof that can contribute to the distribution of voiced geminate and repaired segments is required.

References


CONCATENATIVE AFFIXATION IN BRAZILIAN PORTUGUESE TRUNCATED FORMS*

Ana Paula Scher
Universidade de São Paulo

1. Introduction

Data like those in (1) to (3) have given rise to a fruitful debate in morphophonological frameworks for the study of natural languages (Ito, 1990; Mester, 1990, Weeda, 1992):

1) a. commy-commo       b. communist         Australian English (AE)
2) a. Repro            b. reproduction (Reproduktion) German (G)
3) a. secreta          b. secretária (secretary)     Brazilian Portuguese (BP)

Defining the nature of the word-formation processes that generate these nominal forms has enriched the debate, and their supposedly unpredictable structural properties have led to the suggestion that they do not involve concatenative affixation and to the development of phonological analyses for the phenomenon.

Some of the interesting points which called researchers' attention can be summarized by the two questions below:

a) What is the nature of the word-formation process that generates truncated forms (TFs)?

b) What is their structural representation?

 Apart from these questions, others related to word structure (e.g. Can TFs’ be attributed a predictable structural analysis?) and anchoring (e.g. Which part of the base survives in the truncated form?) dominated the debate and gave rise to proposals based on a globalist framework for linguistic analysis. In that line, McCarthy & Prince (1986; 1988) claimed that the process of morphological truncation maps the base melody segments to a prosodically defined template. A similar analysis comes from Benua (1995). The author suggests that morphologically truncated words show irregular phonology in order to preserve identity with their source word bases. Finally, Gonçalves (2011), dealing specifically with Brazilian Portuguese (BP) data, observes that TFs in that language preserve the onset of the first syllabic foot (from right to left) of their corresponding full forms. For that author, the process that generates TFs occurs in the domain of prosody.

* Most of the ideas for this paper were developed during my sabbatical at University College London. I thank the faculty members at the Department of Phonetics and Linguistics and the LLSD participants to whom I had the chance to talk about this research. Special thanks go to Andrew Nevins, John Harris and Kevin Tang. I also thank the members of GREMD/USP who have always been willing to hear me talk about these truncated forms.

1 Fundação de Amparo à Pesquisa do Estado de São Paulo, Grant: 2011/10818-5.
This paper focuses on two types of TFs in BP, mainly on their structure and their colloquial and evaluative interpretation. Originally, the idea of this paper was to deal with all types of TFs found in BP trying to offer a uniform analysis for all of them. However, the observation of the facts has suggested that this uniformity in the analysis may not be possible due to the different characteristics exhibited by the different TFs. The two types of TFs selected to be discussed in this paper share some structural properties and a special evaluative interpretation, which we claim to be the result of the presence of an evaluative morpheme, namely, \([\text{EVAL}]\) in their morphosyntactic structure.

The organization of the paper is as follows: in section 2, some examples of truncated nominal forms in BP will be presented and grouped into four different types, from which two will be investigated in this paper. Section 3 presents the essentials of the Distributed Morphology (DM) approach to grammar which supports the analysis to be suggested here. The analysis for TFs itself will be presented in section 4, contemplating the two special types of TFs in BP under investigation in this paper, and will account for their special evaluative interpretation in terms of an evaluative head in syntax. Section 5 ends the paper with some final remarks.

2. Truncated words in Brazilian Portuguese

TFs have recently been given more specific attention in the literature on word formation in BP and can roughly be described as in (4):

\[
\text{Truncated Nominal Form (TF)}
\]

corresponds to a compound or derived (suffixed) word types II, III, IV)

(prefixed) word (Class I)

last segment of the root is a consonant (type III, IV)

segment insertion vowel -a (type III) -as, -(i)s (type IV)

Each type of TF is detailed and exemplified below:

Type I: TFs of type I correspond to compound or derived words and are formed by the initial morpheme in these words – a base or a prefix:

5)  a. \textit{psico}\textsuperscript{2} (psicologia) psychology
b. \textit{odonto} (odontologia) dentistry
c. \textit{fono} (fonoaudiologia) speech therapy
d. \textit{super} (supermercado) supermarket

Type II: TFs of type II correspond to primitive or derived words from which the whole root or part of it is preserved. It can be observed that the last segment of the TF is a vowel belonging to the root of its corresponding full form:

6)  a. \textit{deprê} (depressão/deprimido) depression/depressed
b. \textit{pregui} (preguiça/preguiçoso) laziness/lazy

\textsuperscript{2} The stressed syllable is in bold.
Type III: TFs of type III correspond to primitive or derived words from which the whole root or part of it is preserved. They differ from TFs of type II in that the last segment of the root is a consonant and the TF is formed by the insertion of a segment: the vowel –a is added. Type III is the most common type of TF in BP:

7) a. *furta* (fortuna) fortune
   b. *neura* (neurose) neurosis
   c. *secreta* (secretária) secretary
   d. *delega* (delegado) sheriff
   e. *gurja* (gorgeta) tip

These three more basic types of TFs in BP have already been described and discussed in the literature. They have received different analysis in terms of prosodic morphology and OT (Belchor, 2005, 2006, 2009, Gonçalves, 2006, 2009, 2011, Gonçalves & Vasquez, 2004), as well as in terms of Distributed Morphology (Scher, 2011, 2012a, 2012b).

Scher (*op. cit.*) observed the emergence of truncated nominal forms ending in a sequence such as –as or –is, as described below:

Type IV: TFs of type IV correspond to primitive or derived words from which the whole root or part of it is preserved. Like TFs of type III, they differ from TFs of type II in that the last segment of the root is a consonant. When compared to TFs of type III, however, the difference between the two types rests in the segment inserted after the root to form the TF: those of type IV are formed by the insertion of one of the two sequences of segments, –as or –(i)s.

8) a. *bermas* (bermuda) shorts
   b. *saudas* (saudades) homesickness
   c. *bob(i)s* (bobear) silliness
   d. *brinc(i)s* (brincadeira) game, joke
   e. *vont(i)s* (vontade) will (n.)

As mentioned above, this paper focuses, specifically, on two types of TFs in BP, namely, types III and IV. The reason for this selection comes form the similarities between the two types. It was observed that they are both rhizotonic. In addition, neither of them exhibits any morphemes, except for the root and the thematic suffix. Finally, they differ from those in types I and II in that they add extra material to the truncated form, namely, the vowel -a or the sequences –as or –is.

Two other properties are worth mentioning and have to be emphasized, since they raise important questions about TFs. The first one relates to TFs in BP, in general. Most often, these truncated words are attributed an appreciative reading by the speakers of BP. This special reading has already been observed and described in the literature (Gonçalves, 2005; Villalva, 2008), but it still lacks a formal explanation. Therefore, the question that remains unanswered is the following: how does this special interpretation obtain? Can it be related to the structure of these forms? The second property relates to TFs of types III and IV, specifically and has to do with the difference between TFs and their corresponding full forms. It has been observed that the former (TFs) lack some phonological segments which appear in

---

3 The letter “c” is pronounced as /k/, as in *brincadeira* (game, joke), *brinquedo* (toy), etc.
the latter (full forms) as exponents for word categorizing abstract morphemes. For example, the forms -ist- and -a appear in the word *baterista* as the phonological realization of the syntactic N morpheme and the thematic nominal suffix, respectively. The word *batera*, on the other hand, does not exhibit the phonological material for the realization of the abstract syntactic N morpheme, but it seems to realize the nominal thematic suffix, with the theme vowel -a.

When it comes to nominal thematic suffixes in BP and their corresponding theme vowels, it is worth mentioning Alcântara (2010), who, following the work of Harris (1999), presents a DM based description and analysis for groups of non derived BP non-verbal words ended in the non stressed vowels /o, a, e/ or in phonological “zero”. In her paper, the author suggests that a post-syntactic operation of the type “add morpheme”, motivated by an idiosyncratic requirement similar to a condition on word formation rules, adds to the word structure a terminal node for the nominal thematic suffix, which confers the structure the status of an independent word. This morphological terminal node will be phonologically realized by the non stressed vowels /o, a, e/. The structural representations she suggests for the syntactic and morphological derivations of nominal forms in BP can be seen in (9):

9) a) Syntax
   X
   b) morphology
   X
   X  3

Having said that, the observation of the data from (10) to (12) suggests that some morphological properties of BP TFs can, in fact, be predicted. When compared to their corresponding full forms, TFs of types III and IV can be said to be missing the phonological realization for one or more derivational suffixes and, also, for the non-verbal form class morphemes or nominal thematic suffix, when applicable:

10) a. *chin-a*   b. *chin-ês* (Chinese)   suffix
11) a. *bater-a*   b. *bater-ist-a* (drummer)   suffix + nominal thematic suffix
12) a. *forast-a*   b. *forast-eir-o* (outsider)   suffix + nominal thematic suffix

Similar behavior can be observed in the examples from (13) to (16). The difference between these examples and those from (10) to (12) relates to the number and type of suggested abstract morphemes without phonological realization in the TFs:

13) a. *deleg-a*   b. *deleg-a-do* (sheriff)   verbal thematic suffix + suffix + verbal thematic suffix
14) a. *adren-a*   b. *adren-al-in-a* (adrenalin)   suffix + suffix + nominal thematic suffix
15) a. *jap-a*   b. *jap-on(tä)o-ês* (japanese)   suffix + suffix
16) a. *brinc-(i)s*   b. *brinc-a-(d)eir-a* (joke)   verbal thematic suffix + suffix + nominal thematic suffix

Based on the data from (10) to (16), we formulate the following generalization for the TFs of types III and IV in BP:

Generalization: **TFs grouped in types III and IV in BP, when compared to their corresponding full forms, miss phonological material that has morphological status, standing for abstract morphemes in the full forms.**
Previous analyses for the phenomenon, in terms of phonological theory, do not always offer an account for either the generalization above or the appreciative reading observed for TFs. Concerning the generalization, the foot analysis offered by Gonçalves (2011), for example, suggests that TFs are formed from their corresponding full forms, through a phonological process that drops part of the last foot of the word, keeping its onset. Such an analysis clearly does not account for data such as *adrena*, the TF corresponding to *adrenalina* (adrenalin), since it predicts that the sequence of segments *-ina* should be dropped from the foot *-lina*. The base for the formation of the TF would then be *adrenal*, deriving *adrenala*, contrary to facts. Most importantly, the previous analyses do not consider the fact that, most often, the sequence of phonological segments, missing in TFs, has morphological status. As for the appreciative reading attributed to TFs, most researches mention it but do not provide a formal account for it.

Further investigation suggested that TFs in examples from (17) to (19) might, in principle, contradict the generalization formulated above, since the sequences of phonological segments that they are missing cannot be said to be canonical affixes in their corresponding full words.

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>17</td>
<td>a. cerv-a</td>
<td>b. cerv-eja (beer)</td>
</tr>
<tr>
<td>18</td>
<td>a. pij-a</td>
<td>b. pij-am (pijamas)</td>
</tr>
<tr>
<td>19</td>
<td>a. burocr-a</td>
<td>b. burocr-ata (bureaucrat)</td>
</tr>
</tbody>
</table>

The forms –ej-, in *cerveja, –am- in pijama or –at- in burocrata are not supposed to be considered separate morphemic units in these words. (17c), (18c) and (19c) are tentative representations for the words *cerveja, pijama and burocrata*, which go back to their Gallic or French origins. These representations, however, may not be available for present day speakers, who, according to the tentative analysis I’ll present here, might be treating –ej-, -am- and –at- as derivational suffixes. Therefore, they find it easy to segment the words *cerveja, pijama and burocrata* as in (20), treating cerv-, pij- and burocr- as their roots.

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>a. √cerv-ej-a</td>
<td>b. √pij-am-a</td>
</tr>
</tbody>
</table>

Thus, in spite of not being canonical morphemes in the words mentioned above, the forms –ej-, –am- and –at-, among many others, are being treated as such by the speaker. Evidence for this claim comes from two facts. First of all, the forms –ej-, -am- and –at- occur as suffixes in other contexts. Thus, the form –ej- is a real diminutive suffix and occurs as a real suffix in examples such as those from (21) to (24):

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>21</td>
<td>broto-ej-a</td>
<td>(rash)</td>
</tr>
<tr>
<td>22</td>
<td>sertan-ej-a/o</td>
<td>(a woman/man from “sertão” - backlands)</td>
</tr>
<tr>
<td>23</td>
<td>pardal-ej-a/o</td>
<td>(small sparrow)</td>
</tr>
<tr>
<td>24</td>
<td>grac-ej-o</td>
<td>(joke)</td>
</tr>
</tbody>
</table>

Likewise, the form –am- is a collective suffix and occurs as a real derivational suffix in examples such as those from (25) to (28):

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>dinheir-am-a</td>
<td>(lots of money)</td>
</tr>
<tr>
<td>26</td>
<td>poeir-am-a</td>
<td>(lots of dust)</td>
</tr>
<tr>
<td>27</td>
<td>cabel-am-a</td>
<td>(lots of hair)</td>
</tr>
<tr>
<td>28</td>
<td>burac-am-a</td>
<td>(lots of holes)</td>
</tr>
</tbody>
</table>

Finally, the form –at- is a real nominalizing suffix, as can be seen in examples such as those from (29) to (32):
To argue in favor of the morphological status of -ej-, -am- and -at- in the contexts shown in (20), it is worth observing that in those contexts these sequences of segments attract the stress of words in which they appear. According to Camara Jr (1970), a derived word in Portuguese will never have a stressed root: roots in derived words systematically transfer their underlying stress to the last suffix in the word. This can be seen in all examples from (21) to (32), in which, the relevant attested suffixes are stressed. As mentioned above, all full forms in (20) place their stress within the last sequence of segments and, therefore, the speaker will see this as a sign to treat them as derived words, in which -ej-, -am- and -at- are analysed as suffixes. TFs may, then, be formed without some parts of these derived words, as long as they are morphemic units.

The same reasoning can be made for examples such as those from (33) to (37):

<table>
<thead>
<tr>
<th>Example</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>33) a.</td>
<td>fort-un-a</td>
</tr>
<tr>
<td></td>
<td>(fortune)</td>
</tr>
<tr>
<td>b.</td>
<td>furta</td>
</tr>
<tr>
<td>34) a.</td>
<td>vagab-und-a</td>
</tr>
<tr>
<td></td>
<td>(=slut)</td>
</tr>
<tr>
<td>b.</td>
<td>vagab-a</td>
</tr>
<tr>
<td>35) a.</td>
<td>vestib-ul-ar</td>
</tr>
<tr>
<td></td>
<td>(=entrance examination at university)</td>
</tr>
<tr>
<td>b.</td>
<td>vestib-a</td>
</tr>
<tr>
<td>36) a.</td>
<td>gorg-et-a</td>
</tr>
<tr>
<td></td>
<td>(=tip)</td>
</tr>
<tr>
<td>b.</td>
<td>gurja</td>
</tr>
<tr>
<td>37) a.</td>
<td>analf-ab-et-o</td>
</tr>
<tr>
<td></td>
<td>(=illiterate)</td>
</tr>
<tr>
<td>b.</td>
<td>analf-a</td>
</tr>
</tbody>
</table>

When accepting and producing furta, vagaba, vestiba, gurja or analfa as the truncated forms for fortuna, vagabunda, gorgeta and analfabeto, respectively, the speaker seems to be reanalyzing the mono-morphemic words fortuna, vagabunda, gorgeta e analfabeto as the bi-morphemic or even multimorphemic forms fort-un-a, vagab-und-a, vestib-ul-ar, gorg-et-a and analf-ab-et-o. Thus, in spite of not being a morpheme in the words above, the forms realized as -un-, -und-, -ul-, -ar, -et-, -ab- and -et-, among many others in different contexts, are being treated as such by the speaker.

Evidence for this comes from the phonological properties of these sequences of segments, which behave as real suffixes of the language, since they attract the stress of the derived words in which they appear. As a matter of fact, all full forms in (20) and those from (33) to (37) place their stress within the last sequence of segments and will, therefore, be treated as derived words, from which some morphemic units might be absent in the derivation of a truncated corresponding form.

3. Distributed Morphology

Halle & Marantz (1993) developed the Distributed Morphology approach to grammar, a model of linguistics analysis in which the grammar has no lexical generative component. According to this model, both words and sentences are formed under the rules of the syntactic component. One of the consequences of this proposal is that the analyses based on the lexicon in previous models have now to be discussed in terms of formal, semantic and phonological properties which are distributed all over the distinct components of the grammar. The architecture of the grammar includes a Morphological Structure (MS) which serves as the syntax-phonology interface.
There are three central properties defining this model:

a) **Late insertion** – syntactic categories are purely abstract, with no phonological content. The phonological expression of syntactic terminal nodes is inserted only in the mapping to PF;

b) **Underspecification of Vocabulary Items (VIs)** – the phonological expressions do not need to be completely specified for the syntactic positions where they will be inserted in. Default VIs can be inserted when an more specified item is not available;

c) **Hierarchical syntactic structure all the way down** – elements in syntax and morphology undergo the same types of syntactic operations.

Apart from that, some post-syntactic morphological operations are proposed to solve the mismatches between the organization of the terminal nodes at the syntactic level, on one hand, and at MS and PF, on the other. These morphological operations apply to the syntactic structures, before they are filled with phonological expressions. Terminal nodes may be added to the structure, they may be moved from one point to another, merged, fissioned or fused. They may even have some of the features they bring from the syntax deleted in certain contexts.

The analysis for the TFs to be developed in this paper will explicit and explore some of these operations. In particular, the addition of terminal nodes in the morphological component will play an important role in the proposal.

4. **An analysis for TFs of types III and IV**

This paper argues in favor of a more morphological account for the so-called TFs of types III and IV, which involves rescuing a concatenative analysis to the word formation process that derives these types of TFs.

Under this approach, TFs of types III and IV can be treated as derivations from the root, rather than the result of any kind of process of segment deletion applied to their corresponding full forms. They are independently derived by a syntactic process which involves root categorization and the concatenation of an evaluative head to the structure, followed by a morphological operation which inserts a thematic suffix in the derivation, because of an idiosyncratic condition on word formation in BP (as argued by Alcântara, 2010). Subsequently, rules of vocabulary insertion apply and vocabulary items (VIs) are placed on the root and other terminal node positions, such as category morphemes and thematic suffixes, the former, inserted in the syntactic derivation, and the latter, inserted in the morphological component.

The representations below illustrate this idea. The diagrams in (39a,b) and (40a,b) show both the syntactic and morphological structures for the full form *baterista* (drummer) and its corresponding TF *batera* (drummer), respectively.
The Proceedings of GLOW in Asia IX

39) a) syntactic derivation

\[
\begin{array}{c}
\sqrt{\text{BATER}} \\
\text{N}
\end{array}
\]

b) morphological derivation

\[
\begin{array}{c}
\sqrt{\text{BATER}} \\
\text{N}
\end{array}
\]

(39a) represents the syntactic step of the categorization of the root in the formation of the word *baterista*. At the morphological structure, as can be seen in (39b), a terminal node for the thematic suffix (\(\exists\)) is added to the derivation. No further operation modifies the representation of the full forms and vocabulary insertion applies as in (39b).

The syntactic and morphological derivations of the TF of the types III and IV, in (40), differ from the derivation of their corresponding full forms. In (40), the derivation of TFs of the types III and IV involve the presence of [EVAL], an evaluative category which accounts for their appreciative reading.

40) a) syntactic derivation

\[
\begin{array}{c}
\sqrt{\text{BATER}} \\
\text{N}
\end{array}
\]

b) morphological derivation

\[
\begin{array}{c}
\sqrt{\text{BATER}} \\
\text{N}
\end{array}
\]

[EVAL] does not substitute for the N head which categorizes the root and defines the meaning of the word, as we saw in (39a). It only adds the appreciative reading to the structure which will generate the TF, keeping, for the derivation in (40a), the same category and roughly the same meaning as those in the derivation in (39a). At the morphological structure in (40b), before vocabulary insertion applies, a rule of impoverishment (Bonet, 1991) such as (41) deletes the feature [n] in the category N in the presence of the head [EVAL]:

\[
N[n] \rightarrow \emptyset / [EVAL].
\]

For that reason, when it comes to vocabulary insertion, no VI marked with the [n] feature, such as –*ist*– (42a), for example, will be an adequate alternative for insertion in the terminal node corresponding to N, since it has a feature which is no longer present in the structure. No VI will, then, be inserted in the N head. [EVAL], on its turn, will be filled by a null VI as in (42b). At the morphological structure, terminal nodes for thematic suffixes (\(\exists\) ) are added to the derivation and are filled with Vocabulary Items such as those in (42c) (Alcântara, 2010).

42) a) [-ist-] \(\rightarrow\) [n]

b) \(\emptyset \rightarrow [\text{EVAL}]\)

c) [-o] \(\rightarrow\) [Class I], [-a] \(\rightarrow\) [Class II], [-e] \(\rightarrow\) [Class III]
Evidence for the presence of the EVAL head in the structure of TFs of types III and IV comes from examples such as those from (43) to (45), which are considered as truncated words by some speakers, but are, in fact, cases of words with an appreciative reading that can be attributed to one of its morphological units:

43) a. pad-oc-a \( \text{(padaria)} \) baker shop  
44) a. feij-uc-a \( \text{(feijoada)} \) Brazilian typical dish made with black beans  
45) a. metr-anc-a \( \text{(metralhadora)} \) riffle

The examples above exhibit explicit marks of diminutive – -oc-, -uc- and -anc – but no diminutive meanings. Rather, the meaning expressed by the words from (43) to (45) converge to an evaluative reading. Their syntactic and morphological derivation can be seen in the representations in (46a,b) below:

46) a) syntactic derivation       b) morphological derivation
\[
\begin{array}{l}
\text{N} \\
\text{Eval} \\
\text{[Class II]} \\
\sqrt{\text{PAD}} \quad \text{N} \\
\sqrt{\text{FEIJ}} \quad \text{N} \\
\sqrt{\text{METR}} \quad \text{N} \\
\sqrt{\text{ROOT}} \\
\text{|} \\
\text{|} \\
\text{|} \\
\text{|} \\
\text{|} \\
\text{|} \\
\text{|} \\
\text{|} \\
\text{|} \\
\text{|} \\
\text{\( \emptyset \)} \\
\text{\( \emptyset \)} \\
\text{\( \emptyset \)} \\
\text{-oc-} \\
\text{-uc-} \\
\text{-anc-} \\
\text{-a} \\
\text{-a} \\
\text{-a}
\end{array}
\]

The analysis suggested here highlights the morphological properties of the formation of TFs of types III and IV in BP, characterizing this process as a concatenative one just like the one which derives their corresponding full forms.

5. Final Remarks

In this paper, I argued in favor of a morphological approach to, at least, two types of truncated forms in BP: those which correspond to derived words and preserve the root of their corresponding full forms, adding to them the vowel -a or one the sequences of segments, -as or (i)s.

My claim is that these forms are, in fact, derived independently from their corresponding full forms through a concatenative process of word formation. For that reason, it is not correct to assume that there is a truncation process to derive them: they are syntactically derived from the root, to which a categorizing nominal head and an evaluative morpheme are associated. The presence of the evaluative morpheme leads to the impoverishment of the n feature in the N head, which, consequently, prevents this node to be realized by an ordinary n-categorizer. No phonological expression will, then, be able to realize the derivational suffix in the structure of the TF. They are formed by the root, plus a categorizing and an evaluative head in syntax, as well as a thematic suffix, which is added in the morphological component, in observation of the idiosyncratic condition on word formation in BP, identified in Alcântara (2010).
References


McCarthy, J. & A. Prince (1986) *Prosodic Morphology*, ms., University of Massachusetts, Amherst, and Brandeis University, Waltham, Massachusetts.


AUDIENCE DESIGN AFFECTS CLASSIFIER POSITIONING IN CHINESE RELATIVE CLAUSE

Yanan Sheng and Fuyun Wu
Shanghai International Studies University

1. Introduction

There are two prominent models in the field of sentence production: the Audience Design Model and the Speaker-internal Constraints Model. The Audience Design Model (Clark & Murphy, 1982; Temperley, 2003) posits that speakers tailor their utterances to the audience for the ease of comprehension, using ambiguity-avoidance devices or providing cues for difficult structures. Evidence for the Audience Design Model mainly comes from studies on the choice of referential expressions (Arnold, Eisenband, Brown-Schmidt, & Trueswell, 2000; Arnold & Griffin, 2007). These studies find that speakers use gender-marked pronouns more often when there are two characters of different genders than when there are two characters of the same gender, suggesting that speakers choose referential expressions accordingly to help hearers identify the referent efficiently.

In contrast, the Speaker-internal Constraint Model (Lindblom, 1990; Ferreira & Dell, 2000) proposes that speakers produce more accessible words or syntactic structure first, in order to minimize their production efforts and buy time to plan more difficult linguistic materials (Race & MacDonald, 2003; Arnold, Wasow, Asudeh, & Alrenga, 2004). Evidence for the Speaker-internal Constraint Model mainly comes from availability effects in a wide range of structures involving order of mention (Bock, 1982, 1986, 1987). Recently, Ferreira & Dell (2000) investigated the use of optional complementizers “that” in English relative clauses (RCs). They failed to find ambiguity-avoidance effects, as speakers did not show the tendency to insert “that” in ambiguous reduced RCs. Instead, they found that speakers tended to omit the optional “that” if the linguistic information after “that” was mentioned before, thus lending support to the availability hypothesis.

Both the Audience Design Model and the Speaker-internal Constraints Model have supporting evidence from English, thus more research is needed from other languages. Chinese relative clause construction in conjunction with the demonstrative classifier phrase stands as a good test case to distinguish these two models, as detailed below.

1.1. Flexible Ordering of Demonstrative-Classifiers in Chinese Relative Clauses

Chinese RC construction is typologically unique, because it combines the (S)V0 word order and noun-final properties (Greenberg, 1966; Dryer, 1994). Chinese RCs are head final (or pre-nominal), with the head noun following the RC (1), whereas English RCs are head initial (or post-nominal), with the head noun preceding the RC (2).

(1) head-final RC in Chinese

\[ \text{[RC xingtian yujiand e] nusheai} \\
\text{star-hunter meet DE girl} \]

(2) head-initial RC in English

the girl [RC that star-hunter met]
Other than the head-final properties, Chinese RC does not have relative pronouns as in English. The relative clause marker, DE, appears at the end of the clause, and is located right before the head noun, as illustrated in (1).

With these typologically distinctive properties, Chinese RC has become a focus of research among psycholinguists in recent years. Existing work on Chinese RC processing has mostly focused on two most common types of RC cross-linguistically (cf. Keenan & Comrie, 1977), namely, subject-extracted RCs (SRCs) and object-extracted RCs (ORCs). In SRCs (3), head nouns are extracted from the subject position; in ORCs (4), head nouns are extracted from the object position.

(3) Subject-extracted RCs

\[\text{RC} \ yujian \ xingtan \ de \ nuhai \ zhangde \ hen \ piaoliang\]

meet \ star-hunter \ DE \ girl \ look \ very \ beautiful

‘The girl who met the star-hunter looks very beautiful.’

(4) Object-extracted RCs

\[\text{RC} \ xingtan \ yujian \ de \ nuhai \ zhangde \ hen \ piaoliang\]

star-hunter \ meet \ DE \ girl \ look \ very \ beautiful

‘The girl whom the star-hunter met looks very beautiful.’

Moreover, Chinese is known as an isolating language with little morphological inflections. One exception is its numeral classifier system, in which a classifier must agree with its modified noun phrase. Related to the RC construction, the classifier-noun agreement implies that a classifier must be semantically congruent with its head noun.

Crucially related to the present research, when a demonstrative-classifier (DCL) sequence co-occurs with a RC, the DCL can occur either before or after the RC, forming two constructions: pre-RC classifier construction (5) in SRC, or post-RC classifier construction (6) in ORC.

(5) Pre-RC DCL in SRC

\[\text{na-ge} \ \text{RC} \ yujian \ xingtan \ de \ nuhai \ zhangde \ hen \ piaoliang\]

that-CL \ meet \ star-hunter \ DE \ girl \ look \ very \ beautiful

‘The girl who met the star-hunter looks very beautiful.’

(6) Post-RC DCL in ORC

\[\text{RC} \ xingtan \ yujian \ de \ nuhai \ zhangde \ hen \ piaoliang\]

start-hunter \ meet \ DE \ that-CL \ girl \ look \ very \ beautiful

‘The girl whom the star-hunter met looks very beautiful.’

The flexible word order of DCL in Chinese RCs stands as a good case to detect the two models, because they make different predictions regarding the distribution pattern of DCL positioning in RCs.

1.2. Contrastive Predictions Made by the Two Models

When applied to Chinese sentences (5-6) as above, the Audience Design Model and the Speaker-internal Constraints Model make distinct predictions on DCL positioning in RCs of different extraction types. Let’s first examine the predictions made by the Speaker-internal Constraints model, which posits that speakers tend to produce whatever is easiest first in order to plan more difficult linguistic units. According to Hawkins’ (1983) Heaviness Serialization
Principle\textsuperscript{1}, RCs are longer or heavier in syntactic weight than demonstratives or numerals. Thus, DCLs should be easier than complex RC structure, and be available first for speakers to produce, regardless of the extraction type of RC. Therefore, the Speaker-internal Constraints Model predicts no differences in DCL positioning between (5) and (6).

Now we turn to the Audience Design Model, which posits that speakers tend to tailor their utterances to the audience for the ease of comprehension. They may use ambiguity-avoidance devices or other strategies to help hearers overcome comprehension difficulties brought by certain syntactic structures. It argues that speakers organize their utterances in ways that can provide their listeners with more cues in order to help listeners better understand their words and phrases. Thus the Audience Design Model predicts that speakers would help listeners to pre-build RC structure by providing early cues, which in this case are the DCLs. Therefore, speakers tend to use pre-RC DCLs.

However, in ORCs, pre-RC DCLs might cause structural ambiguity resulted from a classifier congruent with both local and head nouns, as in the case of (7), or semantic clash incurred by the incongruence between a mismatching classifier and a local noun, as in the case of (8).

(7) Pre-RC DCL in ORC (structural ambiguity)
\begin{verbatim}
na-ge [Rel xingtan yujian de] nuchai zhangde hen piaoliang
that-CL human star-hunter meet DE girl look very beautiful
\end{verbatim}
‘The girl whom the star-hunter met looks very beautiful.

(8) Pre-RC DCL in ORC (semantic clash)
\begin{verbatim}
na-kuai [Rel baoan duokai de] shikuai diaozai dishang.
that-CL stone/*human guard dodge DE stone fall ground
\end{verbatim}
‘The stone that the security man dodged fell to the ground.

Taken together, the Audience Design Model predicts an asymmetric pattern of DCL positioning: DCLs tend to occur at the left edge of SRCs (5), but at the right edge of ORCs (6), in order to avoid potential ambiguity (7) or semantic clash (8).

1.3. Review of Existing Work

Existing work on the frequency of different types of RCs in Chinese or the distribution of DCLs in Chinese RCs has almost exclusively focused on written corpora (Tang, 2007; Pu, 2008; Wu, 2009; Ming & Chen, 2010). A unanimous finding was that SRCs occur more frequently than ORCs, and that classifiers tend to occur at the left edge of SRCs (i.e. in pre-RC positions), but at the right edge of ORCs (i.e., in post-RC positions) (Tang, 2007; Pu, 2008; Wu, 2009; Wu, Kaiser, & Andersen, 2010; Ming & Chen, 2010).

However, very few studies examined spoken corpus, to our best knowledge perhaps only two — Tang (2007) and Pu (2008), each with a small portion of spoken corpus data. Apparently, more research needs to be done to clarify whether the pattern found in written corpora can be extended to spoken corpora.

Furthermore, few of related studies have investigated the distribution of classifiers in RCs. Existing few studies exploring this issue suggest that mismatching classifiers may induce great lexical disruption in accessing the local noun (Hsu, Phillips, & Yoshida, 2005), but given facilitative discourse contexts, they could provide cues for ORC structure-building (Wu, Haskell, & Andersen, 2006; Hsu, 2006; Wu, Kaiser, & Andersen, 2009). Evidence from the Chinese Treebank 5.0 corpus data, however, shows that there is virtually no token of

\textsuperscript{1} Specifically, the Heaviness Serialization Principle states as follows:
\begin{center}
Rel > Gen > Adj > Dem/Num ( “>” means longer or heavier than)
\end{center}
classifier mismatch-match configuration in ORCs in natural language use (Wu, 2011), suggesting that the potential lexical disruption effect induced by a mismatch between the classifier and the adjacent noun might be deliberately avoided in the written news corpus.

However, it remains unknown whether the finding of Wu’s news corpus study can be generalized to oral speech because, presumably under great time pressure, speakers may be less listener-oriented or altruistic than writers or editors in avoiding potential lexical disruption (e.g., Mims & Trueswell, 1999; Ferreira & Dell, 2000).

In sum, previous studies on the correlation between DCL positioning and RC-extraction types were either conducted on rather small oral data sets (Tang, 2007; Pu, 2008) or exclusively on written data (Wu, 2009; Ming & Chen, 2010). To fill the research gap, we conducted a spoken corpus investigation and two word-based sentence production experiments.

1.4. Organization of the Paper

The present study aimed to investigate the distributional pattern of DCL positioning in Chinese RCs and its underlying reasons from the perspective of sentence production. Specifically, we use both spoken corpora and behavioral data to (i) further our understanding of how native speakers of Mandarin Chinese process RCs, and to (ii) evaluate and refine currently available accounts of language production models in general.

The rest of the paper is organized as follows. In Section 2, we present a spoken corpus investigation of classifier positioning in different RC extractions, and offer possible reasons underlying the distribution pattern from the perspective of language production strategies. In Section 3, we report two experimental studies to verify the distribution pattern obtained from the spoken corpus and to further test the influence of animacy in the DCL-RC processing. In Section 4, we summarize the major findings, and discuss the implications of the spoken corpus study and the experimental study.

2. A Spoken Corpus Study

The purpose of the spoken corpus analysis was to examine the distribution pattern of DCL positioning in RCs, and to see whether the pattern found fits the predictions of the two models.

2.1. Method

The spoken corpus consisted of five volumes of transcribed texts (totaling 309,848 words) from a Chinese live TV-show similar to Oprah Winfrey’s talk show — 

Lu Yu You Yue

(‘Appointment with Luyu’).

A total of 356 relative clauses were extracted, and manually coded for the following factors: (1) presence vs. absence of the demonstrative + Classifier (DCL) sequence; (2) head noun’s syntactic role in the main clause (5 types); (3) extraction types of RC (subject- or object-extracted); (4) position of DCL sequence (pre-RC or post-RC); (5) RC-internal verb types (transitive, intransitive, ditransitive).

2.2. Results and Discussion

Out of 356 RCs, only 128 (or 35.96%) have DCL sequences, in contrast to 228 tokens (or 64.04%) that do not contain classifiers. Given our interests, the discussions below focus on those 128 target RCs.

Taking into consideration of RC extraction types, out of 128 DCL-RC co-occurring tokens, 100 (78.13%) are SRCs, and 28 (21.88%) are ORCs. The difference is statistically
significant \( (p < .0001) \). This result replicated the pattern reported in previous corpus studies, written or spoken, that SRCs occur more frequently than ORCs (Hsiao & Gibson, 2003; Tang, 2007; Pu, 2008; Wu, 2009; Ming & Chen, 2010).

Figure 1 shows the distribution pattern of DCL positioning in SRCs and ORCs. As we can see from Figure 1, there is an asymmetric distribution of classifiers as a function of RC extraction types: Out of 100 SRCs, 85 (85%) occurred in pre-RC positions and 15 (15%) in post-RC positions. The difference is statistically significant \( (p < .0001) \). In contrast, out of 28 ORCs, 20 (71.43%) occurred in post-RC positions, whereas only 8 (28.57%) in pre-RC positions. The difference is again statistically significant \( (p < 0.05) \). Thus, SRCs prefer to occur in pre-RC positions, and ORCs in post-RC positions.

Figure 1 Distribution pattern of DCL positioning in SRCs and ORCs

The most important finding in the current corpus study is that the distribution pattern clearly demonstrates that SRCs tend to have pre-RC classifiers, whereas ORCs tend to have post-RC classifiers in Chinese. This asymmetric classifier positioning pattern in SRCs and ORCs suggests that speakers in the TV show were considerate enough, designing their utterance by taking into consideration of listeners. Particularly in ORCs, speakers deliberately avoided pre-RC classifier to ease the lexical retrieval difficulty for their listeners. This pattern is predicted by the Audience Design Model, but not by the Speaker-internal Constraints Model.

It is worth noting that the asymmetrical distribution pattern of DCL positioning in SRCs and ORCs found in the current spoken corpus replicates the result of previous written corpus studies (Tang, 2007; Pu, 2008; Wu, 2009; Ming & Chen, 2010). The strikingly similar distribution pattern found in corpora of different sources and genres sheds light on the common processing mechanism underlying the human language.

Although the spoken corpus analysis proves to be a useful tool to uncover the correlation between DCL positioning and RC-extraction types in natural utterance production, the conclusion may still not be readily extended to a general population due to the small size of the current corpus. It is also possible that preparations before live TV interviews may have influenced the final product of the current spoken corpus in ways that do not occur in natural setting.

To partly address these concerns, two real-time experiments were designed using oral sentence production tasks.

3. Two Sentence Production Experiments

Two on-line sentence production experiments were conducted to determine whether the results of the corpus analysis could be extended to oral production. Experiment 1 manipulated
the RC types, keeping both head and embedded nouns animate. To eliminate the potential ambiguity incurred by the double animate configuration, Experiment 2 crossed RC types with contrastive animacy configuration (i.e. inanimate head, animate embedded noun). The findings of these two experiments directly tested the predictions made by the two language production models.

3.1. Experiment 1

3.1.1. Method

3.1.1.1. Participants

Forty-two students from a University in Shanghai participated in the experiment. All were native speakers of Mandarin Chinese and they were paid for their participation.

3.1.1.2. Materials and Design

Twenty-four sets of critical sentences were constructed. Each consisted of four components: a DCL sequence (e.g., na-ge ‘that-CL’), an RC (e.g., yujian xingtan de ‘meet star-hunter DE’), a head noun (HN) (e.g., nuehai ‘girl’) and a main clause (MC) (e.g., zhangde hen piaoliang ‘grow very beautiful’). Each set had two RC-type conditions (SRC vs. ORC), as shown in (9a-b).

(9) a.  SRC condition
na-ge     yujian  xingtan   de nuehai zhangde  hen  piaoliang
that-CL    animate meet star-hunter DE girl looks very pretty
‘The girl who met the star-hunter looks very beautiful.’

b.  ORC condition
xingtan   yujian  de  na-ge     nuehai zhangde  hen  piaoliang
star-hunter meet DE that-CLanimate girl looks very pretty
‘The girl whom the star-hunter met looks very beautiful.’

In addition to the target items, there were 48 filler items. Similar to target items, all fillers were chunked into four components, and their presentation layouts were counterbalanced.

The four components of each sentence were assigned to four rectangular boxes on a visual display (Figure 2). The positions of boxes were fixed on the screen: top, left, right, and bottom. To randomly assign key words into the rectangular boxes—specifically, DCL, HN and RC, we kept the distance between DCL and RC the same as the distance between DCL and HN. Such visual equidistance could minimize the potential effects on mental conceptualization of an utterance from the physical distance between words shown on the display. This resulted in 8 possible versions of visual display.
Figure 2 Visual layouts with sample stimuli

However, a DCL in the top box would result in a head noun and an RC right below it, left or right. Thus, the linear presentation of words might lead a “visual” participant to simply read them out, namely either DCL-RC (e.g., ‘that-CL meet star-hunter DE’) or DCL-head (e.g., ‘that girl’) sequences. To eliminate positioning confounds and to make the task more challenging, these two versions were excluded, leaving 6 visual layouts in total for further counterbalancing.

3.1.1.3. Procedure

The stimuli are presented to the participants via Paradigm software on the computer screen. The task used in this experiment was similar to that used in Ferreira (1996) and Huang and Kaiser (2011). Participants sat in front of a computer screen and used the space bar on the keyboard to initiate and end trials. The sentences to be produced in each trial were displayed on a laptop, using Paradigm software (Perception Research Systems). Participants’ spoken responses were recorded with a head-mounted microphone. Each trial began with words presented in four rectangular boxes on the screen. The procedure is like this: the four word chunks were shown on the screen for 4500ms, then they disappeared; at the same time, participants heard a warning tone. They could start to speak out a sentence; once they finished, they pressed the spare bar to proceed to the next trial. If they did not say anything within 20 seconds, the program would go to next trial automatically. It took about 20 minutes to complete the whole experiment.

3.1.2. Responses and Data Coding

Forty-two participants yielded a total number of 1008 responses, including 504 responses for subject-gap condition and 504 responses for object-gap condition. In ORC condition, four null responses were discarded, and due to a script error, ten responses were removed. Thus, 504 responses in SRC condition and 490 responses in ORC condition were included in statistical analyses.

To ensure reliability in transcription, the data recorded were carefully transcribed and were entered into a spreadsheet for coding and analysis. Sometimes participants made two responses for one single trial, however, only the second was reported in the results below because the participants were allowed to correct themselves as long as the time permitted.
An utterance was coded into one of the following three categories:

1. **Perfect** (as expected). The utterance matched the target sentence, without any new or changed words.
2. **Unexpected** (but grammatical). The utterance contained addition, deletion or alternation of words (specifically, two NPs), but still contained an RC with unexpected “exchange errors”.
3. **Wrong**. The utterance was a phrase instead of a sentence, typically a determiner phrase (DP). The utterance also counted as ‘wrong’ if it was not well-formed and semantically unacceptable.

### 3.1.3. Results and Discussion

Given the focus of this study, we only discuss the first category, the perfect responses, because those utterances contain the target structure. There are more perfect (as expected) utterances in the SRC condition than the ORC condition (88.1% vs. 84.7%). In 444 target SRCs, participants overwhelmingly prefer to use pre-RC DCL and were less likely to use post-RC DCL (77.5% vs. 22.5%; \( p < .0001 \)).

In the case of ORCs, we find no preference, though numerically there are a few more post-RC DCLs than pre-RC DCLs (51% vs. 49%). The distribution preference of pre-RC DCL and post-RC DCL in SRC and ORC condition is not clear.

However, if we shift our perspective slightly by observing the distribution percentage of SRCs and ORCs in pre-RC DCL and post-RC DCL sentences, we get an asymmetric distribution pattern: in 548 pre-RC DCLs, there are much more SRCs than ORCs (62.7% vs. 37.3%; \( p < .0001 \)), while in 311 post-RC DCLs, there are more ORCs than SRCs (68.1% vs. 31.9%; \( p < .0001 \)). This asymmetry appears to support the Audience Design.

As discussed above, in Experiment 1 we found the expected asymmetric pattern of DCL positioning in SRC condition. However, the asymmetry appears not that distinct in the ORC condition, wherein no significant differences were found between pre-RC and post-RC DCLs. This may be due to the double animacy configuration used in Experiment 1, which rarely occur in natural language corpora (Hsiao & Gibson, 2003; Kuo & Vasishth, 2006; Wu, 2011). It is further noted that RCs with two animate NPs may potentially induce similarity-based interference (Gordon, 2001; 2004). Moreover, Wu, Kaiser, & Andersen (2012) found a late facilitatory effect of contrastive animacy configuration in on-line study comprehension study.

To investigate whether animacy affects DCL positioning in ORCs, Experiment 2 used a contrastive animacy configuration, namely, animate EN and inanimate HN.

### 3.2. Experiment 2

Experiment 2 was designed to verify the results of Experiment 1 and to eliminate the confounding effects might be caused by the double animacy configuration. Predictions are similar to these of Experiment 1. In addition, the Audience Design Model predicts that animacy affects DCL positioning in ORCs, because the lexical disruptions can be avoided by putting DCLs in the post-RC position. However, the Speaker-internal Constraints Model predicts no effects of animacy.

### 3.2.1. Method

#### 3.2.1.1. Participants

Forty-eight students from a university in Shanghai participated in the experiment. All were native speakers of Mandarin Chinese and they were paid for their participation. They did
not participate in Experiment 1.

3.2.1.2. Materials and Design

The materials were constructed in the same way as in Experiment 1, except that in Experiment 2, all the RC HN were inanimate, and the DCL only could modify the RC inanimate HN, but not the RC embedded noun.

The target sets of sentence were mainly adapted from Wu et al. (2012), in which two NPs and RC-verbs are generally well controlled in terms of frequency and plausibility. Some words were modified or changed. Moreover, specific classifiers (e.g., kuai) were used for corresponding head nouns. One target example was given below:

(10)  a. SRC condition
   na-kuai  zazhong  baoan  de shikuai diaozai dishang
       that-Cl        stone       guard       DE stone    fell to ground
   ‘The stone which hit the security man fell to the ground.’

       b. ORC condition
   nakuai   baoan duokai de shikuai diaozai dishang
       that-CL      stone       dodge      DE stone    fell to ground
   ‘The stone which the security man dodged fell to the ground.’

3.2.1.3. Procedure

The procedure was the same as in Experiment 1.

3.2.2. Responses and Data Coding

Forty-eight participants yielded a total number of 1152 responses, including 576 responses for subject-gap condition and 576 responses for object-gap condition. In SRC condition, four null responses were discarded. After trimming, 572 responses in SRC condition and 576 responses in ORC condition were included in the statistical analysis. As in Experiment 1, the data were coded in three categories: Perfect (as expected); Unexpected (but grammatical); Wrong.

3.2.3. Results and Discussion

Again, we only discuss the perfect responses, as those utterances contain the target structure. In 515 target SRCs, participants overwhelmingly preferred to use pre-RC DCL, and they were less likely to use post-RC DCL (71.5% vs. 28.5%; \( p < .0001 \)); in 460 ORCs, they preferred to use post-RC DCL, and were less likely to use pre-RC DCL (33.9% vs. 66.1%; \( p < .0001 \)).

Moreover, if we shift our perspective slightly, observing the distribution percentage of SRCs and ORCs from the perspective of DCL positioning, again we get the asymmetric distribution pattern: Out of 523 pre-RC DCLs, there were much more SRCs than ORCs (70.2% vs. 29.8%; \( p < .0001 \)), but out of 452 post-RC DCLs, there were more ORCs than SRCs (67.4% vs. 32.6%; \( p < .0001 \)). This result lends further support for the Audience Design Model.

Thus in Exp. 2, we found the strong asymmetrical pattern both within RC-extraction conditions and within DCL positioning conditions. The pattern obtained from Exp. 2 is clearer than Exp.1, in which strong preference was only shown within DCL positioning conditions. The difference between the Exp. 1 and Exp. 2 confirmed the influence of HN-EN animacy configuration.
The experimental results of Exp. 2 also confirmed the results of previous written and spoken corpus studies, that is, there has a connection between classifier positioning and RC-extraction types. Moreover, the distribution preference found in the Exp. 2 replicates that of the spoken corpus study as we reported in section 2.

3.3. General Discussion

To summarize, we have two key findings in on-line sentence production experiments. First, there is a correlation between RC-extraction type and DCL positioning: SRCs prefer pre-RC DCL, ORCs prefer post-RC DCL. Second, as shown in Exp.2, animacy configuration modulates DCL positioning.

The pattern got from this experiment clearly demonstrates that there is a significant correlation between DCL positioning and RC-extraction types. In SRCs, the participants prefer to use pre-RC DCL, but in ORCs, they prefer post-RC DCL.

This significant correlation further confirms the Audience Design accounts of language production. According to Audience Design Model, speakers tend to produce pre-RC DCL in SRCs. On one hand, the pre-RC DCL in SRCs can give the speakers more time to construct the upcoming RC structure. On the other hand, pre-RC DCL in SRCs also can help listeners build the upcoming RC structure.

However, speakers would prefer to use post-RC DCL in ORCs. This is because the pre-RC DCL in the ORC condition might bring to listeners great lexical disruption (Hsu et al., 2005) or semantic clash (Wu, 2009) in processing the mismatched HN. Thus speakers preferred to use post-RC DCL to help listeners avoid the possible lexical disruption or semantic clash. Taken together, the distribution pattern of DCL in SRCs and ORCs in Exp. 2 further supports the Audience Design Model.

4. Conclusion

In this study, we presented the results of a spoken corpus study and two word-based sentence production experiments designed to investigate two different issues related to the DCL-RC structure in Mandarin Chinese – DCL distribution pattern and language production models. The most important findings from the spoken corpus study and experiments are summarized in the following:

First, consistent with the earlier findings in the written corpora, DCL positions are asymmetrically correlated to RC types in the present study. That is, pre-RC DCL prefers to occur in SRCs, and post-RC DCL prefers ORCs.

Second, this correlation has been consistently replicated in the two on-line sentences production experiments.

Third, animacy configuration modulates DCL positioning in online RC productions in Mandarin Chinese.

In conclusion, the corpus and behavioral data can be consistently accounted for by the Audience Design Model of language production. That is, speakers tailor their utterance to help their audiences to anticipate intended structure, and to avoid potential ambiguity or lexical disruption in complex structures. The findings of this research will advance our understanding of the Chinese RC processing and the general mechanism that underlies language production.
Reference


1. Introduction

A common view in the phonological literature is that phonological alternations typically involve markedness reduction; that is, deviation from an input is basically characterized as a repair of a marked structure. Under the most extreme version of phonological theory, every output constraint is simply a reflection of a markedness hierarchy, meaning that all disparities between inputs and outputs are attributed to markedness reduction (see de Lacy 2007 for relevant discussion; cf. Moreton 1999). While the jury is still out as to whether all phonological alternations should be explained in terms of markedness reduction, it is generally agreed that markedness is the major driving force behind phonological alternations.

However, it is reported that in certain cases, phonological alternations seem to derive otherwise disfavored forms (Benua 1995; Tessier 2010). This paper presents novel data from Japanese hypocoristics which show some peculiar properties that cannot simply be attributed to markedness reduction. We report the results of two experiments designed to shed light on these properties. The implications of the present study are two-fold: first, the constraint *#D is psychologically real albeit seemingly no longer operative in Modern Japanese; second, Japanese speakers can evaluate markedness relative to a particular stratum of the lexicon, which lends support to the lexical stratification hypothesis put forward by Itô and Mester (1995a, 1995b, 1999).

2. Hypocoristic Formation in Japanese

For starters, let us take a brief look at the basic properties of Japanese hypocoristic formation. It is by now a widely accepted view that hypocoristic formation in Japanese is subject to the so-called bimoraic requirement (see Poser 1984, 1990; Mester 1990). The most common way of forming a hypocoristic in Japanese is by truncation to the first two moras of the name and suffixation of a diminutive (including zero affixation), as illustrated below.

(1) a. /mija.zaki/ + /san/ → mija-san
   b. /jama.guti/ + /tti/ → jama-tti
   c. /isi.gaki/ + /ii/ → iss-ii
   d. /naka.zono/ + /Ø/ → naka

While the pattern exemplified in (1) is of the most common type, there is also a less common
but attested type of hypocoristic formation in Japanese, which is the main topic of this paper. Hypocoristics of this type involve truncation to two moras of the base in non-word-initial position and suffixation of a diminutive (including zero affixation). Thus, the very same combination of a family name and a suffix can give rise to different hypocoristics, as demonstrated by (1), which we just witnessed, and (2) below.

(2) a. /mija.zaki/ + /san/ → zaki-san
   b. /jama.guti/ + /tti/ → guti-tti
   c. /isi.gaki/ + /ii/ → gakk-ii
   d. /naka.zono/ + /Ø/ → zono

However, non-word-initial truncation is not always possible. For instance, all of the following examples are unacceptable to the present authors.

(3) a. /mija.hara/ + /san/ → *hara-san
   b. /jama.kawa/ + /tti/ → *kawa-tti
   c. /isi.tani/ + /ii/ → *tani-i
   d. /naka.mura/ + /Ø/ → *mura

What is worth noting here is that the derived forms in (2) have a word-initial voiced obstruent, whereas the unacceptable hypocoristics in (3) have either a voiceless obstruent or a sonorant word-initially. To a first approximation, we may thus say that hypocoristics with a word-initial voiced obstruent are preferred over those without. However, this is rather odd in light of the fact that word-initial voiced obstruents are considered to be disfavored in native Japanese phonology. Historically, word-initial voiced obstruents were absent in Old Japanese (Vance 1983, among others), which can be captured if we assume that the constraint *#D (Don’t have a word-initial voiced obstruent) was ranked high in Old Japanese. One might argue that this constraint is no longer operative in Modern Japanese (cf. Kuroda 2008), as there are Sino-Japanese words, loanwords, and onomatopoeic words that begin with a voiced obstruent. However, native words with initial voiced obstruents are vanishingly rare even in Modern Japanese. Furthermore, #D is phonetically marked for aerodynamic reasons (Westbury & Keating 1984). Despite these facts, #D seems to be preferred when it comes to hypocoristics of the non-word-initial truncation type. It thus appears that phonologically marked forms are preferred over unmarked ones. Henceforth we call this phenomenon “Deliberate Markedness.”

We have seen that hypocoristic formation in Japanese exceptionally tolerates otherwise illicit forms (see Benua 1995 and Tessier 2010 for similar phenomena in other languages). Note, however, that this does not mean that phonologically marked forms are simply preferred over unmarked forms, nor that hypocoristics can be formed without any restrictions. Just because a hypocoristic has a word-initial voiced obstruent, it does not necessarily mean that it is well-formed, as the following illicit example demonstrates.

(4) /naga.tani/ + /san/ → *gata-san

This suggests that although hypocoristics seem to be exempt from certain phonotactic restrictions, it is not just that constraints are flagrantly violated without any principles. In the example just given above, morpheme-left-anchoring seems to play an important role. Similarly, changing a voiceless obstruent into its voiced counterpart does not result in a well-formed hypocoristic even though the derived form would have a word-initial voiced obstruent.

---

3 Pejorative words and function words constitute a class of exceptions to this generalization.
Deliberate Markedness in Japanese Hypocoristics (Y. Tanaka and J. Yashima)

(5)  a. /kawa.kami/ + /san/ → *#2 gami-san
    b. /kawa.kami/ + /san/ → *#2 gawa-san

This indicates that IDENT-BaseTruncatum[voice] (henceforth IDENT-BT[voice]) along lines with Benua (1995) is ranked high.

To summarize the observations so far, Japanese hypocoristic formation comes in two varieties. The two types of hypocoristic formation differ in that one type involves word-initial truncation, while the other involves word-medial truncation. The latter type displays a rather peculiar behavior; seemingly phonotactically disfavored forms are derived via truncation. In the remainder of this paper, we investigate (i) whether Deliberate Markedness is a systematically governed phenomenon, (ii) why marked forms are preferred over unmarked ones, and (iii) why the constraint *#D is selectively violated. As for (i), we report the results of two experiments, showing that the phenomenon is so systematically constrained that it cannot be dismissed as a collection of minor exceptions, and for (ii) and (iii), we suggest two analytical options, one based on distinctiveness and the other based on recoverability.

3. Experiment 1

To confirm that Deliberate Markedness is not just a hodgepodge of random exceptions but rather is constrained in a systematic manner, we conducted a naturalness rating experiment.

3.1. Participants

Seventeen native speakers of Tokyo Japanese (ten female and seven male; mean age: 27.6) participated in this experiment. They received a gift card with a value of JPY 400 for participation.

3.2. Materials

The stimuli used in this experiment were hypocoristics formed from seventy-two attested family names (all having a CVCV.CVCV structure) chosen out of the two thousand most common family names in an online corpus (see References). The selection of the base names was based on the [voice] and [sonorant] feature values of the consonants involved. From each family name, three to five hypocoristics were created with the suffixation of "-sa-san" in accordance with eight target conditions to be spelled out shortly. Examples are given in (6).

(6) Examples of stimuli

<table>
<thead>
<tr>
<th>Base</th>
<th>Hypocoristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>mija.zaki:</td>
<td>1. mija-san, 2. jaza-san, 3. zaki-san, 4. saki-san</td>
</tr>
<tr>
<td>taka.kura:</td>
<td>1. taka-san, 2. daka-san, 3. kaku-san, 4. kura-san, 5. gura-san</td>
</tr>
<tr>
<td>naɡa.tani:</td>
<td>1. naɡa-san, 2. qata-san, 3. tani-san, 4. dani-san</td>
</tr>
</tbody>
</table>

3.3. Procedure

Participants were given a spreadsheet which contains seventy-two distinct family names

---

4 Some hypocoristics derived by voicing change or morpheme-medial truncation had two voiced segments or two labial segments (e.g. /siba.hara/ → ziba-san, /mizu.quti/ → zuqu-san, /hama.saki/ → bama-san). Since the low acceptability of these hypocoristics might have stemmed from Lyman’s Law (Ito & Mester 2003) or a dispreference for labial-labial sequences (Kawahara, Ono & Sudo 2006), they were excluded from the analysis.
and their corresponding hypocoristics (one hypocoristic per family name) with the order of presentation randomized across subjects. They were asked to rate the naturalness of each hypocoristic on a scale of 1 to 4: 4 = “natural,” 3 = “somewhat natural,” 2 = “somewhat unnatural” and 1 = “unnatural.”

Each family name was presented together with an attested first name. The combinations of the first and last names were randomized across subjects. For each family name, only one hypocoristic was provided to each participant, and no participant saw different hypocoristics derived from the same name. They were asked to read out loud the last name, the first name and the hypocoristic before rating each hypocoristic.

3.4. Conditions and Hypotheses

The design had eight target conditions, which are listed in (7) along with exemplifying base names and hypocoristics.

(7) Target conditions of Experiment 1 $^5$

a. Non-initial truncation deriving #D (e.g. /jama.zaki/ → zaki-san)
b. Non-initial truncation deriving #T (e.g. /hama.saki/ → saki-san)
c. Initial truncation of the same base names as (a) (e.g. /jama.zaki/ → jama-san)
d. Initial truncation of the same base names as (b) (e.g. /hama.saki/ → hama-san)
e. Non-ini. trunc. deriving #D by voicing change (e.g. /kawa.kami/ → gami-san)
f. Initial trunc. deriving #D by voicing change (e.g. /kawa.kami/ → gawa-san)
g. Morpheme-medial truncation deriving #D (e.g. /naga.tani/ → gata-san)
h. Morpheme-medial truncation deriving #T (e.g. /naka.tani/ → kata-san)

Of particular interest to our discussion here is the comparison between Conditions (a) and (b); if Deliberate Markedness is actually at work, hypocoristics in Condition (a) should be rated higher than those in Condition (b). Conditions (c) and (d) are meant to form a baseline because they are the most common type of hypocoristic and are expected to be rated relatively high.

In addition, we also expect that items involving voicing change will be rated relatively low; if we are right in extrapolating from a limited set of data (see the examples in (5)) that IDENT-BT[voice] is ranked high, hypocoristics in Condition (e) should be rated lower than those in condition (a) even though they both have #D. Likewise, hypocoristics in Condition (f) should also receive lower ratings than those in Conditions (c) and (d) alike. Lastly, items in Conditions (g) and (h) involve morpheme-medial truncation, and thus are expected to be rated relatively low if our earlier observation in example (4) derives from a general edge-anchoring requirement.

3.5 Results

Figure 1 plots the average naturalness ratings of hypocoristic formation in the eight conditions, starting with the one that received the highest score. Each condition is accompanied by an illustrative example.

---

$^5$ For simplicity, ‘non-word-initial truncation’ is referred to as ‘non-initial truncation’, and ‘word-initial truncation’ as ‘initial truncation’ unless indicated otherwise. The design also had a condition of non-initial truncation deriving #T by voicing change (e.g. /jama.zaki/ → saki-san). Since this condition was not our main focus here, we treated it as a filler condition.
Deliberate Markedness in Japanese Hypocoristics (Y. Tanaka and J. Yashima)

Figure 1: Experiment 1 results - Average naturalness ratings in eight conditions (error bars: 95% confidence intervals)

As can be seen from the figure above, participants rated the hypocoristics in Condition (a) higher than those in Condition (b). A one-way repeated measures ANOVA indicated a significant main effect of Formation Condition on the ratings of hypocoristics \((F(7, 112) = 97.346, p < 0.0001)\). Post-hoc pair-wise comparisons using Bonferroni’s adjustment for multiple comparisons revealed that the average naturalness ratings of the hypocoristics in Condition (a) were significantly higher than those of the hypocoristics in Condition (b) \((p < 0.05)\). This means that in the case of hypocoristics involving non-initial truncation, speakers preferred outputs that violate *#D over those that don’t, which in turn indicates that Deliberate Markedness is indeed a psychologically real phenomenon.

Also, the difference between Conditions (a) and (e) was statistically significant \((p < 0.0001)\). This indicates that non-initial truncation deriving #D is less acceptable when it involves voicing change. Similarly, the hypocoristics in Conditions (c) and (d) were, respectively, rated significantly higher than those in Condition (f) \((ps < 0.0001)\). These results together support the hypothesis that IDENT-BT[voice] is ranked high. The hypocoristics in Conditions (a) and (b) were, respectively, higher than those in Conditions (g) and (h) \((ps < 0.0001)\), indicating that there is also a strong requirement on morpheme-edge-anchoring.

The overall results of the experiment supported our hypotheses. Most importantly, the difference between Condition (a) and Condition (b) indicates that Deliberate Markedness is actually at work.

4. Proposals

The results of Experiment 1 have shown that hypocoristics with an initial voiced obstruent are preferred over those without. In order to account for this peculiar phonological phenomenon, we propose two possible analyses: the distinctiveness account and the recoverability account. In this section, we first outline the two accounts and the predictions they make. We then discuss ways to tease them apart empirically.

4.1. Distinctiveness Account

One way to account for the Deliberate Markedness phenomenon is to claim that it is driven by speakers’ desire to derive hypocoristics that sound distinctive. Since a hypocoristic is used to individualize a person special to the speaker, distinctiveness is a key property. As discussed above, there are very few native Japanese morphemes having an initial voiced
obstruent. In fact, most native Japanese family names begin with a segment other than a voiced obstruent, namely a voiceless obstruent, a sonorant or a vowel. This in turn suggests that names with an initial voiced obstruent sound unusual and as such serve the function of being distinctive. Thus, our first proposal is that speakers intentionally violate the constraint *#D in order to derive distinctive hypocoristics.

It might be worthwhile to mention in passing that deliberate deviations from certain standards or norms have been documented in phonology and other areas of linguistics. Some English shortenings in the 80’s-90’s also run afoul of shortening norms, thereby creating distinctive and catchy names and titles (e.g. rents for parents, za for pizza, blog for weblog, etc.). Likewise, the Uncola, which was coined by a soft-drink company, yields a similar effect by violating the selectional restrictions of the prefix un-. All these cases can be most naturally viewed as a reflection of speakers’ intention to derive distinctive (individualizing) forms. It is therefore quite plausible to claim that the preference for a marked form in Japanese hypocoristic formation discussed in this paper is another instance of distinctiveness-driven phonological phenomena.

This distinctiveness-based explanation predicts that, other things being equal, not only a *#D violation but violations of other ‘minor’ constraints can in principle achieve the same effect. As stated above, although native words do not usually begin with a voiced obstruent, there are Sino-Japanese words, loanwords and onomatopoeic words that violate the constraint *#D. A simple but, in our view, most viable approach to this heterogeneous nature of the Japanese lexicon is to assume that the Japanese phonological grammar has a stratal structure (Itô & Mester 1995a, 1995b, 1999). On this account, *#D can be analyzed as highly ranked only in the native stratum of the grammar. This approach allows us to entertain the possibility that having an initial voiced obstruent is deviant enough from the Japanese phonotactic standards to derive a distinctive hypocoristic, but not as aberrant as outright phonotactic illegality when the entire grammar of Japanese is taken into account. If this reasoning is valid, we expect that just like a *#D violation, flouting other stratum-specific (and to that extent minor) constraints should be effective in creating distinctive hypocoristics, too. That is to say, a violation of constraints such as *#r (Don’t have an initial [r]), which is highly ranked only in the native stratum (Itô & Mester 1995b), should also fit the bill. How this prediction can be tested will be discussed in Section 4.3.

4.2. Recoverability Account

An alternative explanation can be constructed from the perspective of recoverability. That is, there is an avoid ambiguity strategy according to which truncated material must be interpreted as corresponding to the initial segments of its base word if the application of truncation leads to a loss of recoverability. The underlying idea is that a base needs to be sufficiently recoverable from its truncated form. Under this view, we can explain the unacceptable use of “saki-san” as a hypocoristic for “hama.saki” in the following way. Suppose that there is a hypocoristic “saki-san.” Then the possible corresponding base names for this hypocoristic basically take the form “saki.XX” (e.g. saki.guti, etc.) or “XX.saki” (e.g. hama.saki, etc.), both of which satisfy necessary conditions for forming hypocoristics in Japanese. However, this situation has a recoverability problem such that the output form is no longer informative with respect to the original anchoring position of the morpheme “saki.” The association of the hypocoristic “saki-san” with a base name of the form “XX.saki” is thus blocked in order to resolve this issue. It is for this reason that “saki-san” is a preferred hypocoristic for a person named, say, “saki.guti” but not for “hama.saki.”

The major advantage of the recoverability-based analysis is that it provides a lucid explanation for the question of why “zaki-san”, which is derived via non-initial truncation, is exceptionally permitted as a hypocoristic. That is, if *#D is psychologically real, then one can
infer that the truncated element “zaki” cannot be associated with a base name of the form “zaki.XX”. This means that the application of truncation does not lead to a loss of recoverability because “zaki” can only be associated with “XX.zaki.” Consequently, the avoid ambiguity strategy does not kick in, and the use of “zaki-sa” for “XX.zaki” (e.g. jama.zaki, etc.) is not blocked. Under this analysis, the reason that names like “jama.zaki” can undergo both initial and non-initial truncation receives a straightforward explanation. It should also be noted in this connection that forms like “zaki” have undergone sequential voicing also known as Rendaku - a phonological process whereby the initial segment of the second member of a compound becomes voiced. (e.g. jama + sakura → jama.zakura ‘mountain-cherry’, jama + saki → jama.zaki ‘mountain-promontory (family name)’). Since “jama.zaki” is decomposed into “jama” and “zaki,” but “zaki” does not exist in isolation (while “saki” does), it is fairly obvious that the morpheme-initial voiced obstruent of “zaki” has emerged as a consequence of Rendaku. These properties presumably facilitate the inference with respect to the original anchoring position of truncated material.

This explanation predicts that not only an initial voiced obstruent (in particular that derived via Rendaku) but also other morphemes that do not occur in word-initial position should be equally eligible for non-initial truncation. More specifically, allomorphs or morphemes that appear only as the second member of a compound name should be able to undergo non-word-initial truncation, even if the output form does not have an initial voiced obstruent.

For example, wara, which is an allomorph of hara ‘field’, always occurs as the second element of a compound name. In the course of historical sound change, */p/ changed to */ɻ/ and later to /h/ in word-initial and morpheme-initial position (e.g. */para/ > */ɻara/ > /hara/ ‘field’), whereas */p/ changed to */ɻ/ and later to /w/ in intervocalic position (e.g. */kapa/ > */kawa/ ‘river’). Thus, this sound change split */p/ into two different phonemes (/h/ and /w/) that occur in different positions. However, some proper names such as personal names and place names have applied the latter type of sound change to the morpheme-initial */p/ in the second position of a compound, treating it as an intervocalic */p/ (e.g. */huzi.para/ > */huzi.ɻara/ > /huzi.wara/ ‘wisteria-field (family name)’, cf. NOT /huzi.hara/). This yielded an allophonic relationship between [h] and [w] in some morphemes, the former being a realization of /h/ in word-initial position and the latter an allophone of /h/ in the initial position of the second morpheme of a compound.6 If the recoverability account outlined here is correct, it is expected that those names such as “huzi.wara” and “suga.wara” which contain the historical allophone [w] should be able to undergo non-word-initial truncation (e.g. “wara-san”) precisely because associating the allophone [w] with a word-initial position is not possible.

There are also morphemes that always appear as a non-initial element of a compound name. For example, -no/-na, which is a possessive marker, always follows a possessor noun. In a family name, it appears between two morphemes, i.e. XX.no.YY or XX.na.YY, where XX is a possessor of YY. Given that these morphemes do not occur in word-initial position, it is expected under this analysis that names such as “iti.no.se”, “wata.na.be”, etc. should be just as susceptible to non-word-initial truncation (e.g. “no.se-san”, “na.be-san”, etc).

---

6 This allophonic relationship is a remnant of a historical sound change and is not attributed to any productive phonological alternation. Also, the sound change of */p/ into its ultimate form /w/ did not occur across the board, and thus family names of the same origin can have [w] or [h] in the same position as in “suga.wara” and “suga.hara”. Note, however, that this does not undermine the argument because the historically-driven [w] (as in -wara) does not occur in word initial position anyway.
4.3. Testing the Proposals

As we have seen, the distinctiveness account and the recoverability account make different predictions. Our particular concerns here are how to test these predictions and how to determine which explanation captures the Deliberate Markedness phenomenon more accurately.

Let us first look more closely at the empirical predictions of the distinctiveness proposal, according to which not only a *#D violation but also violations of other stratum-specific constraints such as *#r should make non-initial truncation possible. Our job is, then, to find family names that provide an excellent opportunity to test this prediction. There’s the rub, however. Finding a native Japanese name with a morpheme-initial [r] is virtually impossible precisely because of the constraint *#r. There are names like “kura.moti” which have a non-initial [r], but deriving a hypocoristic like “ramo-sa” would violate the morpheme-edge anchoring requirement, and we expect it to be ruled out independently of the constraint *#r. For this reason, it is hard to create experimental stimuli against which to test the distinctiveness account directly.

On the other hand, the predictions made by the recoverability account can presumably be tested with relative ease. Unlike the distinctiveness account, this analysis predicts that forming hypocoristics of the non-initial truncation type does not necessarily require phonotactic violations. As mentioned above, there are names such as “huzi.wara”, “suga.wara”, “iti.no.se”, “wata.na.be”, etc. that contain an allophone or a special morpheme that is not expected to occur in word-initial position. Although these particular allomorphs or morphemes never appear word-initially in names, word-initial [w] and [n] are phonotactically legal and can occur freely in the native stratum. Thus, the two proposals make different predictions with respect to the acceptability of those hypocoristics of the non-initial truncation type that commit no phonotactic violations. To be more specific, the distinctiveness account predicts that such hypocoristics are ruled out, whereas the recoverability account predicts that they are acceptable.

The next section reports a follow-up judgment task designed to test these predictions.

5. Experiment 2

As a follow-up study, we conducted another naturalness judgment experiment, where Japanese speakers judged the naturalness of hypocoristics derived from various family names. The primary objective of this experiment was to test the validity of the two proposals. The secondary objective was to replicate the results of Experiment 1 with a larger number of subjects in a more natural linguistic setting. For this reason, we used auditory stimuli instead of orthographic representations.

5.1 Participants

Forty-one native speakers of Japanese (twenty-six female, fourteen male and one no response; mean age: 28.0) participated in the study. The recruitment was done via e-mail. They received a gift card with a value of JPY 500 for participation.

5.2 Material

Hypocoristics formed from forty existing family names taken from the same corpus as employed in Experiment 1 were used as the stimuli. Thirty base names had a CVCV.CVCV structure and they were chosen on the basis of the [voice] and [sonorant] feature values of the consonants involved. The other ten were names of three to five moras containing elements that never appear word-initially. These names included, among others, “huzi.wara” (wara: an
allomorph of hara ‘field’), “ja.wata” (wata: an allomorph of hata ‘flag’), “iti.no.se” (-no: a possessive marker), and “itu.ka.iti” (-ka: a classifier). A complete list of the base names is given in (8). From each family name, three to five hypocoristics were created with the suffixation of “-sa” in the same manner as in Experiment 1.

(8) List of base names for the stimuli used in Experiment 2

1. huku.sima
2. kasa.hara
3. kuma.kura
4. simo.hira
5. hiko.saka
6. kita.qawa
7. kata.giri
8. saka.qami
9. kaki.zaki
10. taka.gaki
11. hasi.moto
12. take.naka
13. katu.mata
14. hisa.matu
15. hisi.numa
16. kawa.saki
17. haja.kawa
18. kami.tani
19. kuro.sawa
20. sira.tori
21. tani.quti
22. kana.zawa
23. same.zima
24. tori.go
25. sino.duka
26. hira.jama
27. kura.moti
28. tera.nisi
29. sawa.mura
30. sumi.josi
31. wata.na.be
32. huzi.wara
33. matu.no.o
34. sio.no.ja
35. hati.no.he
36. iti.no.se
37. o.wari
38. ja.wata
39. ja.k.a.be
40. itu.ka.iti

To make audio stimuli, the hypocoristics were pronounced by two native speakers of Japanese (one female and one male). They produced each hypocoristic in isolation and the utterances were recorded onto a digital recorder (SONY Linear PCM Recorder PCM-D50) through a microphone (SONY ECM-959DT) and were digitized at the sampling rate of 48 kHz. The sound files were saved in the .wav format. The recording was done in a sound-attenuated room in the Sophia University Phonetics Laboratory, Tokyo, Japan.

5.3 Procedure

The task was an Internet-based naturalness judgment task using audio stimuli powered by XHTML5 and JavaScript codes in combination with Google Docs questionnaire forms. Participants were first asked to go to the website for the task. The first part of the webpage showed a consent form for a human subject experiment and general instructions. They were told that the questionnaire was about the naturalness of nicknames formed from family names. They were then asked to complete one practice trial, which contained one story (one rating task) given in the same format as the actual test trials.

In the test session, they were asked to read twenty short stories, each of which contained a character called by a family name and a possible hypocoristic formed from the family name. In the stories, the names and hypocoristics were displayed as clickable buttons saying [NAME] and [NICKNAME]. Participants were instructed to click on the buttons to hear the sounds of the names and hypocoristics. They heard a male or female voice depending on the context of the story. At the end of each story, they were asked a question of the form “What do you think if [NAME] is called [NICKNAME] as in the story?”. As an answer to the question, they rated the hypocoristic by clicking a radio button on the screen. They provided their judgments on a scale of 1 to 5: 5 = “natural,” 4 = “somewhat natural,” 3 = “neither natural nor unnatural,” 2 = “somewhat unnatural” and 1 = “unnatural.” They were allowed to hear the auditory stimuli as many times as they liked.

A sample image of part of the experiment is given in Figure 2. The actual stimuli were

---

Since some web browsers do not recognize sound files in the .wav format embedded in the XHTML5 <audio> tag used in the experiment, we also prepared the stimuli in the .mp3 format that had been converted from the original .wav files. It was coded so that the browser plays the .mp3 version only when it cannot read the .wav format. Thus, some participants received the stimuli as .mp3 sounds. For the purpose of our experiment, however, we do not believe that the difference in the sound quality caused by the reformat affected the results.
presented in Japanese.

**Figure 2:** A sample image of a rating task in Experiment 2 (translated into English)  
http://www.linguistics.ucla.edu/people/grads/yutanaka/exp hypocs/sample.html

The hypocoristics presented to each participant were randomly chosen from a set of derived hypocoristics. The combinations of the names and the stories were randomized across subjects. The same stimulus was not presented twice to any participant, and no participant saw different hypocoristics derived from the same base name.

### 5.4 Conditions and Hypotheses

Experiment 2 had eight target conditions, shown in (9) along with exemplifying base names and hypocoristics.

(9) Target conditions of Experiment 28
- a. Non-initial truncation deriving #D  
  (e.g. /jama.zaki/ → zaki-san)
- b. Non-initial truncation deriving #T/#R  
  (e.g. /hama.saki/ → saki-san)
- c. Non-initial truncation deriving #wara, etc.  
  (e.g. /huzi.wara/ → wara-san)
- d. Initial truncation of the base names for (a)  
  (e.g. /jama.zaki/ → jama-san)
- e. Initial truncation of the base names for (b)  
  (e.g. /hama.saki/ → hama-san)
- f. Initial truncation of the base names for (c)  
  (e.g. /huzi.wara/ → huzi-san)
- g. Non-ini. trunc. deriving #D by voicing change  
  (e.g. /kawa.kami/ → kami-san)
- h. Initial trunc. deriving #D by voicing change  
  (e.g. /kawa.kami/ → gawa-san)

If Deliberate Markedness operates across the board in non-initial truncation, hypocoristics in Condition (a) are predicted to be rated higher than those in Condition (b). Both the distinctiveness-based and recoverability-based accounts predict that items in Condition (c) should receive higher ratings than those in (b). Hypocoristics in Conditions (g) and (h) arguably fail to satisfy one of the necessary conditions for hypocoristic formation (i.e.

---

8 For the purpose of clarity, we use the same base names and hypocoristics as in Experiment 1 for examples, although they were not actually included in the Experiment 2 stimuli. R represents ‘sonorants’. For Conditions (b) and (e), we put voiceless obstruents and sonorants into the same category as opposed to voiced obstruents, since they often behave alike in terms of voicing in Japanese (see Mester & Itô 1989 for discussion on the application of Lyman’s Law in Rendaku). The stimuli also included morpheme-medial truncation hypocoristics deriving #T or #R (e.g. /naka.tani/ → kata-sans). Since this condition is orthogonal to the purpose of Experiment 2, we treated these words as fillers.
5. Results

Figure 3 plots the average naturalness ratings of hypocoristic formation in the eight conditions, starting with the most highly rated one. Each condition is accompanied by a representative example.

![Figure 3: Experiment 2 results - Average naturalness ratings in eight conditions](image)

A one-way repeated measures ANOVA indicated a significant main effect of Formation Condition on the naturalness ratings ($F(7, 280) = 32.529$, $p < 0.0001$). Post-hoc pair-wise comparisons using Bonferroni’s adjustment were also carried out. The hypocoristics in Condition (e) were rated significantly higher than those in Condition (b) ($p < 0.0001$), confirming the predictions of both accounts. Although it can be seen that the average rating points of Condition (a) were higher than those of Condition (b), the difference did not turn out to be statistically significant ($p = 0.217$). This means that, although there seems to be a tendency for #D to be preferred, the preference was not robust enough to provide a definite conclusion. The difference between Condition (b) and Condition (c) was not significant either ($p = 1.000$), meaning that the effects of special allomorphs/morphemes were not attested in this experiment.

In sum, the results did not offer us unequivocal evidence favoring one hypothesis over the other. The fact that the effects of Deliberate Markedness were not as evident as in Experiment 1 makes it particularly difficult to draw a valid conclusion from this experiment. In what follows, we discuss possible interpretations of these results.
6. Discussions

6.1 Interpretations of the General Results

The results of Experiment 1 demonstrated a systematic preference for #D hypocoristics over hypocoristics with no *#D violations, providing support for the existence of Deliberate Markedness. In Experiment 2, however, the effects of Deliberate Markedness were not salient, at least not to such an extent as to allow us to conclude that Deliberate Markedness strongly operates. Also, the difference between Condition (b) and Condition (c) in Experiment 2 was not significant, contrary to the prediction of the recoverability account. That said, an interesting fact can still be found in the pair-wise comparisons between the non-initial truncation conditions and their initial truncation counterparts. As we just observed, the hypocoristics in Condition (b) were rated significantly lower than those in Condition (e). On the other hand, the average rating scores of Condition (a) and Condition (c) were not significantly lower than those of Condition (d) and Condition (f), respectively (ps = 1.000). This leaves the possibility that non-initial truncation is usually less acceptable than initial truncation, but it ameliorates acceptability if it derives hypocoristics with an initial voiced obstruent or a special allomorph in initial position. Then the reason that the experiment did not yield clear-cut results might possibly be attributed to some confounding factors that are at this stage unknown. Since we cannot simply draw a conclusion from non-significant results, we take these results not as evidence for the recoverability account (let alone for any other claim) but as hints for future research.

Note also that, in both experiments, hypocoristics in non-initial truncation conditions were rated relatively high, sometimes even as high as those in initial truncation conditions. This was rather surprising at least to the present authors in light of the fact that non-initial truncation is of extremely rare occurrence.9 Perhaps, the relatively high acceptability of non-initial truncation might have been caused by the existence of those hypocoristics that are clearly unacceptable, namely those derived by voicing alternation and morpheme-medial truncation. We thus cannot deny the possibility that the ratings of these extremely unacceptable cases affected those of other hypocoristics in general. That is, due to the fact that hypocoristics derived by voicing alternation and morpheme-medial truncation are completely ill-formed, the general rating scores of initial truncation and non-initial truncation might have been raised so much as to blur the differences within the non-initial truncation conditions. If this scenario is what actually happened, the effects that we expected to show were possibly somewhat masked in our experiments.

We should also consider the possibility that the phenomenon is driven by the cumulative effects of distinctiveness and recoverability; it might be that only when the conditions for distinctiveness and recoverability are simultaneously met (as in /jama.zaki/ → zaki-san) do we find the phenomenon. We would like to leave all these issues for future research.

6.2 Implications: Psychological Reality of *#D and the Lexical Strata

While certain properties of Deliberate Markedness are still shrouded in mystery, the results of the present study have some theoretical implications. As has been discussed throughout the paper, native Japanese words disfavor initial voiced obstruents, whereas non-native words tolerate them. One way of capturing this fact within the framework of Optimality Theory is to propose that the constraint *#D is ranked high only in the native

9 Indeed, all of the examples given in Poser (1990) and Mester (1990) are of the initial truncation type, and they do not even deal with non-initial truncation cases.

10 See Tanaka (in prep.) for proposals and experiments along these lines.
stratum and that it is outranked by IDENT[voice] in other strata. To implement this idea, one needs to assume that the Japanese phonological grammar has a stratal structure and that constraints can be reranked at different strata (see Itô & Mester 1995a, 1999, among others).

However, this lexical stratification hypothesis has been challenged by some researchers. For example, Rice (1997) questions Itô, Mester and Padgett’s (1995) argument that *NT (Don’t have a voiceless obstruent after a nasal) is active in the native stratum but not in others. Her claim is that, if voicing is contrastive after a nasal in a fair amount of words in the language (i.e. words of Sino-Japanese origin, mimetics, etc. can have a voiceless obstruent after a nasal), leaners should simply conclude that the contrast exists in the entire grammar. On this view, the distributional differences based on etymology might well be attributed to historical factors, which are less of an issue for theories of synchronic grammar, and one would not even need to posit two different strata. We do not go into the details of this particular discussion on *NT here. Interested readers are referred to Itô, Mester and Padgett’s (1999) reply to Rice. Turning back to our discussion on initial voiced obstruents, one could also argue that *#/D is no longer operative in Modern Japanese (cf. Kuroda 2008), attributing the relative rarity of initial voiced obstruents to historical factors. If this is true, then there seems to be no compelling reason to assume that the Japanese lexicon is stratified into several layers (see Gelbart & Kawahara 2007 for a summary of the relevant discussions on the lexical stratification hypothesis and its competing hypotheses).

As we have seen, however, Japanese hypocoristic formation (of the non-initial truncation type) is sensitive to *#/D. This suggests that Japanese speakers know that initial voiced obstruents are marked at a particular stratum. Some recent experimental studies have shown that lexical stratification is psychologically real in Japanese speakers’ minds (Moreton & Amano 1999; Gelbart & Kawahara 2007). This study can be taken as another piece of evidence for the psychological reality of the stratal organization of the Japanese phonological grammar.

6.3 Further Issues

The phenomenon also raises an issue as to how to implement Deliberate Markedness within a grammar model. If the phenomenon is driven by speakers’ desire to violate *#/D as hypothesized by the distinctiveness-based account, we will need a grammar that allows not only violable constraints but also the ones that need to be violated. In a model assuming weighted constraints, one possible approach to pursue is to posit constraint weights on both sides of zero, although many models limit weights to a single side (see Pater 2009 for relevant discussion). Notice, though, that the question still remains as to how we can account for the fact that violations of certain markedness constraints are favored only in particular contexts.

Similarly, the recoverability account also leaves us with a question of how we can implement the blocking effects as envisaged under this analysis (aside from the fundamental question of how recoverable is recoverable enough). As we have seen, the core idea of this analysis is that if truncation creates an ambiguity, the occurrence of a form that is expected to be well-formed is obviated. Note that one cannot simply postulate a constraint that directly penalizes ambiguity. Such a constraint does not fall under the jurisdiction of phonology because its violation does not depend on the structure of a form per se. In other words, whether or not a given sequence has a recoverability problem cannot be determined simply by looking at its phonological form alone. All in all, whether the key factor is distinctiveness or recoverability (or possibly both), the Deliberate Markedness phenomenon constitutes an important issue to the current theories of phonology.
7 Concluding Remarks

In this paper, we discussed an unusual phonological phenomenon in Japanese hypocoristic formation which we dub Deliberate Markedness, where marked forms seem to be preferred over unmarked forms. We presented two experiments that were designed to shed light on the properties of this phenomenon. The results of the experiments indicated that the phenomenon is governed in a systematic fashion. Admittedly, the present study as it stands does not provide a deep analysis of why this phenomenon should occur at all, and to that extent it is vulnerable to criticism. Nonetheless, it has important implications; namely, our findings provide empirical support for the psychological reality of the constraint *#D and the lexical stratification hypothesis.

References

Tessier, A.-M. (2010) “Short but not sweet: Markedness preferences and reversals in English hypocoristics,” A poster presentation at Canadian Linguistics Association Meeting, Concordia University, Montreal, Quebec.
**GLOW in Asia IX 2012**

**Date:** September 4-6, 2012  
**Place:** Sansui Hall, Mie University, Tsu, Mie, Japan  
**Hosted by:** Mie University;  
                Center for Multicultural Studies,  
                Faculty of Humanities, Law and Economics,  
                Mie University