Proceedings of the
Seventh Glow in Asia Conference
2009

Proceedings of GLOW in ASIA VII 2009

edited by
Rajat Mohanty
Mythili Menon

Preface

The 7th meeting of GLOW in Asia was held at the English and Foreign Languages University (EFL‑U), Hyderabad, 25-27 February 2009. Incidentally, this was the venue of the first Asian GLOW, hosted by the same institution (then called the Central Institute of English and Foreign Languages). The theme of the conference was *Universals and Variation*. The three-day conference was followed by a one-day workshop on the theme *Theoretical Understanding of Language Acquisition*, on 28 February 2009.

The conference had two keynote speakers:

Guglielmo Cinque (Università Ca' Foscari di Venezia)
Veneeta Dayal (Rutgers University)

and eight invited speakers:

Josef Bayer (Universität Konstanz)
George N. Clements (CNRS/Sorbonne Nouvelle, Paris)
Probal Dasgupta (Indian Statistical Institute, Kolkata)
Paul Kiparsky (Stanford University)
Heejeong Ko (Seoul National University)
Mamoru Saito (Nanzan University)
Peter Svenonius (University of Tromsø)
Wei-Tien Dylan Tsai (National Tsing Hua University)

Besides the 10 invited talks, the conference consisted of 29 selected papers, in the areas of Phonology, Syntax, Semantics and Language Acquisition.

The workshop had only invited talks. The speakers were:

Andrea Gualmini (University of Leiden)
Thomas Hun-Tak Lee (Chinese University of Hong Kong)
Keiko Murasugi (Nanzan University)
Colin Phillips (University of Maryland)
Tetsuya Sano (Meiji Gakuin University)
William Snyder (University of Connecticut)
Sharon Unsworth (University of Leiden)

The present volume contains 17 of the papers presented at the conference, which are the ones submitted to us in response to the call for papers for a volume of

conference proceedings. We thank the authors for responding to us and for getting the papers ready in reasonably quick time.

We present this volume in the hope that it will be a useful contribution to the ongoing and exciting enterprise of Linguistics.

K.A. Jayaseelan
December 2009

Acknowledgments

We note with deep sadness the passing away of George N Clements,
an eminent linguist and one of our invited speakers, on August 30 2009

Contents

Five Notes on Correlatives[1]

Guglielmo Cinque
Università Ca' Foscari di Venezia, Italy

Abstract

The article attempts to show that correlatives (at least those that do not contain multiple wh-phrases) are embedded in a DP which is left dislocated at the beginning of the matrix clause and is resumed by a correlative pro-form (or a full DP) inside the matrix clause. The left dislocated DP may contain, depending on the language, either an externally headed postnominal, or an externally headed prenominal, or an internally headed, or a headless (free), relative clause (whence the conclusion that correlatives are not an independent, fifth, type of relative constructions. The article also documents the existence of non-restrictive correlatives, and argues that multiple correlatives are not relative clauses at all.

Introduction

In Cinque (in preparation) (see Cinque 2008 for a preliminary presentation) it is proposed that the different types of relative constructions found across languages (externally headed post-nominal, externally headed pre-nominal, internally headed, 'headless' (or 'free'), correlative, and 'adjoined' or extraposed) derive from one and the same structure, whether they involve a raising or a matching derivation.

This unique structure, in compliance with Antisymmetry (Kayne 1994), has the relative clause merged pre-nominally, in a specifier of the extended projection of the NP; more precisely between the position of numerals (and other weak determiners, in the sense of Milsark 1974), and that of demonstratives (and other strong determiners, like the definite article and universal quantifiers).[2]

---

[1] Parts of this article were presented at the 7th Glow in Asia (Hyderabad, February 25-27, 2009), and the 4th Lissim Summer School (Kausani, Uttarakhand, June 10-30, 2009). I wish to thank the audiences of the two events, in particular R. Amritavalli, Tanmoy Bhattacharya, Probal Dasgupta, Veneeta Dayal, K.A. Jayaseelan, Alice Davison, Richard Kayne, Ghanshyam Sharma, and Alessandro Zucchi for discussing specific points of the analysis with me.

[2] An independent conceptual argument for the prenominal origin of relative clauses appears to come from the pervasive left-right asymmetry of natural languages discussed in Cinque (2009). I take this asymmetry to suggest that the complements, modifiers, and functional heads associated

Guglielmo Cinque

See (1), which represents the (simplified) structure underlying the relative clause *the expensive books that John bought*.

(1)
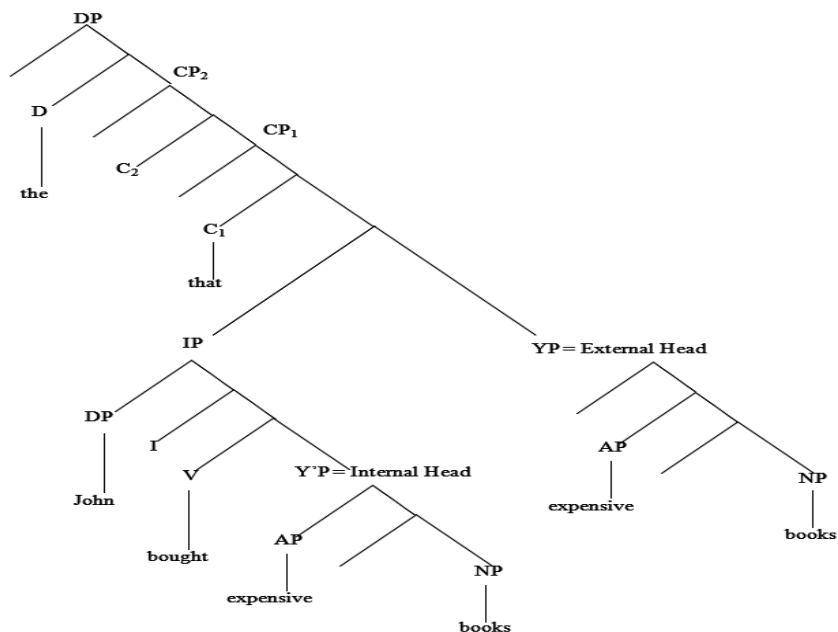
The phrase directly modified by the relative clause (YP in (1)) is the *external Head* of the relative clause, which is matched inside the relative clause by an identical phrase (Y'P, the *internal Head*).

Whenever interpretive factors require reconstruction of the *overt* Head inside the relative clause (idiom chunks, pronominals within the Head bound inside the relative clause, etc.), it is the internal Head which raises to a position c-commanding the external Head (Spec,C_1), causing the latter to delete. Instead, when nothing forces reconstruction of the Head inside the relative clause, the *overt* Head is the external Head, which raises to a position c-commanding the internal Head (Spec,C_2), whether the latter moves or not, and deletes (or 'reduces') it. See Krapova (2009) for evidence to this effect from Bulgarian relatives.

For present purposes it suffices to note that under this analysis all relative constructions, 'headless'/ 'free' relatives included, are double headed (they have

with a lexical head (N, V, etc.) should be merged exclusively to the left of the lexical head, their possible surface location to its right being a function of the raising of a projection of the lexical head to their left. See Cinque (2009) for an elaboration of this point.

both an external and an internal Head). For example, English 'headless'/'free' relative clauses would receive the following analysis, arguably with recoverable deletion (from the particular wh-phrase involved) of such functional nouns as THING, AMOUNT, PLACE, TIME, PERSON,...[3]

(2)   a. (I don't like) [$_{DP}$[$_{CP}$ what THING$_i$ you said t$_i$] (SUCH) THING]

   b. (He weighs) [$_{DP}$[$_{CP}$what AMOUNT$_i$ I weigh t$_i$] (SUCH)AMOUNT]

   c. (Here is) [$_{DP}$[$_{CP}$where PLACE$_i$ they slept t$_i$] THERE PLACE]

   d. (Come)   [$_{DP}$[$_{CP}$when TIME$_i$ you can t$_i$] THEN TIME]

   e. (He helps) [$_{DP}$[$_{CP}$whoever PERSON$_i$ t$_i$ needs it](SUCH) PERSON]

    See Cinque (2008, and in preparation) for discussion of such an analysis. If correct, this proposal prompts a reconsideration of certain aspects of the analysis of correlatives.

NOTE 1: *Simple* correlatives as 'left dislocated' *DPs* resumed IP internally[4]

Following a certain tradition, by 'simple correlatives' I mean those correlatives that contain a single *wh*-phrase, like that in (3):

---

[3] For discussion of some of the other contexts in which the presence of such silent functional nouns can be postulated, see Kayne (2004, 2005a, 2007).

[4] See Rebuschi (1999,68) for the similar idea that the correlative clause may just be "la partie visible d'une véritable relative libre topicalisée.", and especially Gupta (1986, chapter 5), who concludes: "Thus, internal [correlative] and postnominal relative constructions display characteristics of "left dislocated" NPs. These same traits are not evident in extranominal [extraposed] relative sentences" (p.91). Also see Lipták (2004), Dasgupta (2006), Butt, King and Roth (2007, §4.3), and Rebuschi (2009, §3.3). As we see below, the term 'left dislocated DP' corresponds in different languages to different types of 'left dislocation' constructions, while the element resuming the relative in the matrix IP may be represented either by a full DP (see for example (i) below, from Marathi- Renuka Ozarkar, p.c. – which incidentally redresses McCawley's 2004,300 generalization), or by a demonstrative (possibly followed by a head noun), as shown in (3), or by an anaphoric pronoun, which can also be silent, depending on the Case it bears, and the particular language involved.

(i) [jyaa   aattaa-c aalyaa   aahet] **Tyaa laal Dres ghaat-le-lyaa   don Chotyaa mulii**…
     which now-emph come-PAST.FEM be-PRES.PL those red dress wear-PAST.PART.FEM two small/young girls…
      'Those two small girls wearing a red dress who have just arrived…'

In languages that have both demonstratives and special anaphoric correlative pronouns, the two may have different semantic consequences. See Bagchi's (1994) discussion on Bangla.

Sometimes the phrase in the matrix IP which resumes the left peripheral relative is considered as the (external) Head of the relative clause. But this is misleading if the correlative pronoun (phrase) is nothing other than a phrase resuming a 'left dislocated' DP (for multiple correlatives, see NOTE 3 below).

Guglielmo Cinque

(3)　　**jo laRkii** khaRii hai, vo (laRkii) lambii hai　　　(*cf.* Dayal 1996,160)
　　　　which girl standing be-PR, she/that (girl) tall be-PR
　　　　'which girl is standing, she is tall'

An influential analysis of this construction takes the left peripheral relative to be a bare CP, adjoined to the matrix IP, which contains a pronominal (or demonstrative) bound by that CP: See Srivastav (1991), and Dayal (1996).[5]

　　　This analysis is the only conceivable one if both simple correlatives and multiple correlatives (those containing more than one wh-phrase, like (4) (Dayal 1996, 1997)) are taken to represent one and the same construction.

(4)　　　jis laRkii-ne$_i$ jis laRke$_j$-ke saath khelaa, us-ne$_i$ us-ko$_j$ haraayaa
　　　　which girl-ERG which boy with play.PAST, she-ERG he-ACC defeated
　　　　'which girl played with which boy, she defeated him'

Clearly a DP analysis for such cases is out of the question since the correlative CP cannot have two external Heads (cf. Downing 1973,13; Dasgupta 1980,291; Srivastav 1988,148; de Vries 2002,147; Bhatt 2005,9; Anderson 2005,5fn3).[6] Correlatives would thus seem to pose a problem for any unified analysis of relative clauses that takes them to be embedded in a DP.

---

[5] Also see Andrews (1975) and Hale (1976). Among the works that essentially adopt this analysis are Bagchi (1994), Bianchi (1999, chapter 3, section 4.1), de Vries (2002, chapter 5, section 6), Cecchetto, Geraci and Zucchi (2006), Leung (2007c), and various contributions in Lipták (2009). Differently from Srivastav (1991) and Dayal (1996), Bhatt (2003, 2005) argues that the CP is not base-generated as an adjunct to the matrix IP, but is moved there from a position inside the matrix IP adjoined to the correlative pronoun or demonstrative (Mahajan 2000,fn.10 also proposes a movement derivation of the left peripheral relative). In this way, the fact that the relation between the CP and the correlative pronoun or demonstrative in the matrix IP is sensitive to islands can be made to follow. A similar analysis is actually adumbrated in de Vries (2002, 149, fn.49), and Dayal herself (1996, chapter 6, section 2.4) admits that the CP can in certain cases be adjoined to the DP containing the correlative pronoun or demonstrative, and also mentions elsewhere (p.183) that the relation between the two, when they are separated, is subject to island constraints.

[6] In addition to (simple and multiple) correlatives, Hindi has externally headed embedded ((i)a) and extraposed ((i)b) postnominal relative clauses, which share properties setting them apart from (simple and multiple) correlatives (see, among others, Srivastav 1991, Mahajan 2000, McCawley 2004, Leung 2007a,b, Butt, King and Roth 2007, §3). Here I will not be concerned with these other types of relative clauses.

(i) a. vo laRkii **jo khaRii hai** lambii hai　　　(Srivastav 1991,642)
　　　　that girl which standing is tall is
　　b. vo laRkii lambii hai **jo khaRii hai**　　　(Srivastav 1991,642)
　　　　that girl tall is which standing is
　　　　'The girl who is standing is tall'

There is however evidence (discussed in Bhatt 2003, 2005) that multiple and simple correlatives do not constitute a homogeneous construction and thus should not be forced under one and the same analysis that "generalizes to the worst case" (that of multiple correlatives).

Some of this evidence will be recalled in NOTE 3 below, where multiple correlatives will actually be taken to be free adjunct clauses (in Izvorski's 2000 sense), along the lines of Dayal's original analysis.[7]

Here it suffices to observe that simple correlatives like those in (3) contain a 'free' relative which may alternate with an externally headed postnominal relative. Compare (3) with (5):

(5)   vo laRkii jo khaRii hai, vo lambii hai   (*cf.* Dayal 1996,152)
        that girl which standing be-PR, she/that tall be-PR
        'which girl is standing, she is tall'

Taking (3) and (5) together into consideration, and the double headed analysis of 'headless'/'free' relatives given in (2), it becomes possible to interpret (3) as having a silent external Head, as in (6):[8]

(6)   [$_{DP}$VO LARKII [$_{CP}$ jo laRkii khaRii hai]] vo laRkii lambii hai
        THAT GIRL which girl standing be-PR, that GIRL tall be-PR
        'the girl who is standing, that girl is tall'

---

[7] Butt, King and Roth (2007, section 5) also give a non relative clause analysis for multiple correlatives (adjunction to IP) distinct from that for simple correlatives (generation in a specifier of the correlative DP).

[8] Gupta (1986,36fn2) explicitly proposes that a Hindi correlative like (i) derives from an externally headed RC like (ii), with deletion of the external Head (also see Mahajan 2000,215):

(i)   jo laRka: la:l kami:j pahne hai wo   mera: bha:i: hai
          which boy red   shirt wearing is that/he I.gen brother is
           'The boy who is wearing a red shirt is my brother'

(ii)   [[wo laRka:] jo laRka: la:l kami:j pahne hai]] wo   mera: bha:i: hai
           That boy which boy red shirt wearing is that/he I.gen brother is

Junghare (1973) also proposes to derive the Marathi correlative forms in (iii) from a structure essentially like (iv), which however is not acceptable for her. Also see Wali (1982):

(iii) a. **to manus**[jo **Ø**   ith∂ kam k∂rto]   **to manus** ajari ahe
        b. **to**   **Ø** [jo **Ø**   ith∂ kam k∂rto]   **to Ø**   ajari ahe
        c. **to**   **Ø** [jo **Ø**   ith∂ kam k∂rto]   **to manus** ajari ahe
        d. **Ø**   **Ø** [jo **manus** ith∂ kam k∂rto]   **to Ø**   ajari ahe
        e. **Ø**   **Ø** [jo   **Ø** ith∂ kam k∂rto]   **to manus** ajari ahe
        f. **Ø**   **Ø** [**Ø**   **Ø** ith∂ kam k∂rto]   **to manus** ajari ahe
          (that)(man)(which)(man) here work does   that (man) sick is
          'the man who works here is sick'

(iv) **to manus** [**jo manus** ith∂ kam k∂rto] **to manus** ajari ahe (*)

Guglielmo Cinque

Veneeta Dayal (p.c.) tells me that she in fact marginally accepts (7/8), which shows the underlying structure of (3) and (5) on its sleeve, so to speak:[9]

(7)   vo laRkii jo laRkii khaRii hai, vo laRkii lambii hai
        that girl which girl standing be-PR, that girl tall be-PR
        'the girl who is standing, that girl is tall'

The same full structure is apparently acceptable (under the appropriate conditions of emphasis) in two other Indo-Aryan languages: Bundeli ((8)a – Ruchi Jain, p.c.) and Maithili ((8)b, from Singh (1980), according to whom it is "cumbersome, though acceptable"(p.34)):[10]

(8)   a. [**ba moRii** [**jo moRii** ThaRii hε]], **ba moRii** lambii hε
            that girl which girl standing is,   that girl tall is
            'The girl who is standing is tall'

        b. [(**o**) **panc-sab** [**jaah**^{i} **panc-sab-kE**^{n} ham niik jakaan^n janait chalianh^i]$_S$]$_{NP}$
            (the) Panch which Panch-PL-OBJ I good way know.PART BE.PAST.AGR,

            o panc-sab…
            the (same) Panch...
            'The Panch whom I knew very well, the same Panch…'

The 'left dislocated' DP, containing the RC, is matched by a resumptive DP (often pronominal/ demonstrative) in the clause. Depending on the language, the 'left dislocated' DP containing the correlative clause may apparently be either an English-type Left dislocation/Hanging Topic (Kashmiri), or a German-type Contrastive Left Dislocation (German, Bulgarian), or a Romance-type Clitic Left Dislocation (for the "correlatives" of Italian).

        As opposed to the other Indo-Aryan languages, Kashmiri is an (SOV) V-2 language. Its finite verb, in main (and complement) clauses, necessarily occupies the second position, following either the subject or a scene-setting adverb, or a focussed phrase or wh-phrase (Hook and Koul 1996, and especially Bhatt 1999,

---

[9] Alice Davison tells me that (8) was accepted by many speakers she consulted. Wali (2006,289) claims that in Marathi too the left dislocated DP may sometimes surface unreduced. See (v) (Renuka Ozarkar tells me that this is indeed possible if one wants to emphasize 'that particular girl', stressing 'ti' at the beginning of the main clause. Otherwise, it is slightly odd ('?')):
(v) **Ti mulgi** [**ji   mulgi** ghari geli] ti ithe rāhte
     That girl which girl home went that here lives
     'The girl who went home lives here'
[10] The same full structure is instead not readily acceptable in Nepali (Samar Sinha, p.c.).

chapter 4).[11] However, if a left dislocated/hanging topic is present, resumed by a demonstrative or pronominal inside the clause, the finite verb is found in *third* position, with a subject or a focussed/wh-phrase occupying the second position. In other words, the left dislocated/hanging topic phrase does not count as a filler for the "first position".[12]

Now, as Hook and Koul (1996,98) show, a correlative clause too "does not count in the V-2 calculation, with the result that the finite verbal element comes in *third* position". See (9)a, which contrasts minimally with (9)b, characterized by a topicalized Headed postnominal relative (not resumed by a correlative element):

(9)   a. [yus   naphar   raath   aay] bi chus yatshaan temyis samikh-un
           [which person yesterday came] I am wanting him.DAT meet-INF
           'I want to meet the man who came here yesterday'

        b. [temyis naphras yus raath   aav] chus bi yatshaan samikh-un
           [the   person who yesterday came] am I wanting meet-INF
           'I want to meet the man who came here yesterday'

Thus, Kashmiri provides direct evidence that one type of correlative clause can occupy the position of left dislocated/hanging topics, preceding the CP space which contains a fronted phrase (in first position) and the finite verb (in second position).[13]

---

[11] As Richard Kayne reminds me, Kashmiri, as opposed to Germanic V-2 languages, allows multiple wh-fronting, with the consequence that the verb may end up not being in strict second position. It also ends up in third position after a Hanging Topic (see the next footnote), or in the presence of a sentence initial yes/no question marker (Koul 2003, §6.2.1.4). Also see Bhatt (1999, §4.1.2.2).

[12] See for example (i)a-b, from Bhatt (1999,103):
(i) a. Tem dop ki, coon kalam, shiilaayi tshooND su
       he said that, your pen, Sheila   found that
       'He said that as for your pen, it is Sheila who found it'
    b. Coon kalam, su goyi me garyi mashith
       your pen, that gone I home-at forget
       'As for your pen, that (is what) I forgot at home'
Bhatt (1999,103f) gives two arguments for the extra-clausal nature of left dislocated/hanging topics in Kashmiri. The first is that it is possible to insert a parenthetical after them, and the second is that they are "always in the nominative case", whereas the co-referential pronoun in the following clause is in the appropriate Case.

[13] If the left dislocated phrase containing the relative clause in Kashmiri is base generated in the left peripheral position rather than moved there, no reconstruction of the left dislocated DP should be possible, nor should its relation with the correlative element be subject to island constraints. This remains to be checked.

Guglielmo Cinque

Hindi, possibly in addition to an English/Kashmiri-type left dislocation construction (Dwivedi 1994a, section 2.2.2), appears to have a topicalization construction involving movement, possibly similar to Romance Clitic Left Dislocation, modulo the presence of non clitic resumptive DP (either a full DP, or a demonstrative pronoun) (Mahajan 1990; Srivastav 1991; Dwivedi 1994a,b). See, in particular Mahajan (2000,fn.10) and Bhatt (2003) for arguments that the correlative relative acquires its left adjoined position by movement, and Bhatt (2003) for the idea that it starts out together with the correlative pronoun (as seen from the possibility of their making up a constituent), and optionally moves out to a left peripheral position stranding the correlative DP.

We follow this analysis here except for the idea that the RC is internal to a DP which together with the correlative DP forms a "big DP" ([[Head RC] [correlative]), much like the "big DP" taken to underlie French Complex Inversion (Kayne 1972) and Romance Clitic Left Dislocation ([$_{DP}$ DP [$_{D}$Clitic]] – Uriagereka 1995,81).

In Bulgarian, differently from Hindi (and other Indo-Aryan languages), the left dislocated DP of the correlative construction is never found adjoined to the resumptive element (Bhatt 2003,529). Rather, it appears to be base generated in situ and matched by a correlative element which obligatorily moves to the front of the main clause (presumably to Spec,FocusP) (cf. Izvorski 1996,12):

(10)   [Kolkoto pari Maria$_k$ iska], tolkova$_i$ tja$_k$ misli če šte j dam t$_i$
        How much money M. wants, that much she thinks that will her I.give
         'She thinks that I will give her as much money as Maria wants'

This is indicated by the fact that, differently from Hindi (Bhatt 2003, section 3.3.1), the left dislocated DP (in (10)) does not reconstruct, as no Principle C violation is to be observed there.

This appears parallel to the non-connectivity variant of German contrastive Left Dislocation:[14]

Hungarian correlatives, which, as Lipták (2004) shows, do not reconstruct inside the main clause to a position adjoined to the correlative element, nor display sensitivity to islands, also appear (pace her own conclusion) to be Hanging Topics. The two putative differences which according to Lipták (2004, 302) distinguish Hanging Topics from Hungarian correlatives may turn out not to be real. Both correlatives and Hanging Topics seem to be root phenomena and indeed, just as with correlatives, there is in general no more than one Hanging Topic per clause (cf. Postal 1971, 136, fn.17; Cinque 1990,58; although some speakers marginally accept more than one).

[14]Namely to (i)a, where no Case connectivity is present, vs. (i)b:
(i) a. Der Karl, dem will ich vertrauen
     The(Nom) Karl, him(Dat) will I trust

(11)   [Wer das sagen wird] dem will ich vertrauen
        who.NOM that say will that.DAT will I trust
        'I will trust who(ever) says that'

In Italian, the element resuming the "correlative" relative is normally a run-of-the-mill clitic, actually the usual resumptive clitic associated with the Clitic Left Dislocated DP that contains the relative clause (though a demonstrative, itself clitic left dislocated, can resume the correlative relative when this is a hanging topic, as in (12c):

(12)   a. **Qualunque promessa** lui potrà far**ti**, non prender**la** sul serio
           whatever promise he will.be.able.to make to you, not take it seriously
           'Whatever promise he may make to you, do not take it seriously'

        b. **Chi fa cose del genere**, credo Ø non debba essere seguito
           who does such things, I.think not has to be followed
           'I do not think that one should follow someone who does such things'

        c. **Chi ti ha appena telefonato**, **quello lì,** proprio non **lo** sopporto
           Who to you has just telephoned , that there really not him I.can.stand
           'The one who just called you, that one really I cannot stand'

From this perspective, the impossibility of stacking correlatives (Srivastav 1996,175-77; McCawley 2004, section 5; Butt, King and Roth 2007, section 2) should be limited to those containing a left dislocated free relative (as free relatives are also known not to be able to stack – Carlson 1977).[15] It should not extend to those correlatives that contain a left dislocated externally headed (pre- or post-nominal) relative clause, or an internally headed one whose Head has not moved, all of which are known to be able to stack. In the next Note I am actually

---

    b. Dem Karl, dem will ich vertrauen
       The(Dat) Karl, him(Dat) will I trust

[15] Stacking of correlatives is claimed to be possible in other Indo-Aryan languages: Konkani (Almeida 1989,304 - see (i)), and Bhojpuri (Shukla 1981, chapter 19, section 4, p.206 – see (ii)):

(i)   jo   a:j   aila, ja-ka   gʰor na, jace   poise sãdlyat, tya mons-ak pedru adar dita
        who today come, who-dat house not, whose money lost, that man-dat Peter help gives
        'Peter helps the man who has come today, who has no home and whose money is lost'

(ii)   ham jaon pʰal   pa:k-i:,   jaon tu: bec-ba: taon kʰa:-b
        I which fruit ripe-3sg.m.fut, which you sell-2sg.m.fut that eat-1sg.fut
        'I will eat that fruit, which will ripen, which you will sell'

Also see Davison (2009, section 2.2.5) for the apparent possibility of stacking in Sanskrit correlatives. However, given that the impossibility of stacking seems to be a general property of relatives involving raising of the internal Head (free relatives, correlatives with a left peripheral free relative, etc. - Carlson 1977; Grosu 2002), one should determine whether such cases truly involve stacking rather than simple asyndetic coordination (cf. McCawley 2004,306).

Guglielmo Cinque

suggesting that all main types of relative clauses can be left dislocated, and thus enter the correlative construction. To reserve the term 'correlative' just to left dislocated free relatives seems, from this point of view, arbitrarily limiting.

NOTE 2: (Simple) Correlatives as a non independent relative clause type

It is often assumed, in both the typological and generative literature, that correlatives are an entirely separate type of relative clause, but if they are DPs (containing a relative clause) in TopP, resumed by a coindexed resumptive phrase in the matrix IP, then one should expect them to be just a particular manifestation of externally headed postnominal, externally headed prenominal, internally headed, and "headless" (or "free") relative clauses, not an independent, fifth, type.

This indeed seems to be the case as the 'left dislocated' DP can contain, depending on the language, any of the other types of relatives. We have already seen that it can contain an externally headed postnominal relative clause (see (5)), or a "headless"/"free" relative clause (see (3) and the Bulgarian, German, and Italian examples in (10) through (12)). It can also contain an externally headed prenominal relative clause resumed by a coindexed phrase in the matrix IP, as shown by the Sinhala (Indo-Aryan) example in (13):[16]

(13)   [ara [hitagena inna] gaenu lamaya], ee lamaya usa i
        that [standing being] woman child, that child tall is
        'That girl who is standing, that girl is tall.'

Finally, the 'left dislocated' DP can also contain an Internally Headed relative clause resumed by a coindexed phrase in the matrix IP, as in the Wappo example (14), or in the Bambara example (15):[17]

---

[16] I owe this example to Lalith Ananda (p.c.). The phonetic transcription follows the one utilized in Ananda (2008).

Sinhala is generally reported (Bhatt 2003,491; Leung 2007c; Lipták 2009a,10) as not having correlatives (as it does not have embedded postnominal relative clauses with relative pronouns, nor their free relative variant). But, if correlatives are not limited to left dislocated free relatives, this is strictly speaking not true.

Languages with both correlatives and prenominal relative clauses have been claimed (Downing 1978,400) not to exist. But, in addition to the case of Sinhala, Dravidian languages and the language isolate Burushaski also have both correlatives and prenominal relative clauses, even though, differently from Sinhala, for correlatives they utilize a free relative (containing an interrogative adjective/pronoun) resumed by a correlative proform (cf. Lakshmi Bai 1985 for Dravidian, and Tiffou and Patry 1995 for Burushaski).

[17] Cf. Keenan (1985,165). Other languages optionally displaying a left dislocated DP with an Internally Headed relative clause resumed by a phrase in the matrix IP are Arizona Tewa (Gorbett

(14)   [i čhuya t'um-ta]   cephi   šoy'i-khi?
       1SG house buy-PST:DEP 3SG:NOM burn-STAT
       I house bought, that one burned down = 'the house I bought burned
       down'   (Thompson, Park, and Li 2006,117)[18]

(15)   deni mi djolen file,   o (deni) ka djan
       girl which is   standing, that (girl) is tall
       'Which girl is standing, that (girl) is tall'   (Dayal 1996,215fn.15)[19]

NOTE 3: Multiple correlatives as non-relative, free adjunct, CPs

In addition to the possibility for simple, but not for multiple, correlatives to
alternate with externally headed postnominal relatives, there is further evidence
that one should distinguish between two separate constructions: one, a DP
(containing a relative CP), adjoined to the resumptive correlative DP, which it
can strand in its movement to the left-periphery of the matrix IP (as shown in
(16a); the other, a base-generated CP, containing one or more wh-phrases, paired
in the matrix IP with corresponding correlative phrases, as in (16b) (cf. Izvorski
2000. I exemplify with English glosses only):

1977, 272), and, possibly, Italian Sign Language (Branchini and Donati 2009), which also
appears to have externally Headed postnominal relative clauses (also entering a correlative
construction). See Bertone (2006), and Brunelli (2006).

[18] Wappo (a Californian language whose genetic affiliation is unclear - Thompson, Park and Li
2006, xi) also has free relatives resumed by a demonstrative correlative pronoun:

(i)   [te **ita** čo?-me]   **cew** ah   te-k'a   čo:-si?
         3SG where go-DUR:DEP there 1SG:NOM 3SG–COM   go-FUT
          'I'll go wherever s/he goes' (Thompson, Park and Li 2006,123)

Thompson, Park and Li (2006) say that "[t]he demonstrative pronoun seems to be required when
it is *cephi*, the nominative form, but optional when it is *ce*, the accusative form" (p.116).

[19] Bambara (of the Mande branch of Niger-Congo) has both left peripheral Internally Headed
relative clauses resumed by an anaphoric phrase/pronoun ((16)), or Internally Headed relative
clauses in argument position, as in (i), below (in both cases the internal Head is marked by a
following modifier, *mi(n)*). In some varieties it also has externally headed postnominal and
extraposed relative clauses (Bird 1968, Zribi-Hertz and Hanne 1995, and references cited there).

(i)   Tyɛ`  ` be n ye so min ye dyo   (Bird 1968,46)
         man the PRES [I PAST house wh- see] erect
          'The man is building the house that I saw'

Guglielmo Cinque

(16)   a. 'Ram, which CD is on sale, that CD bought'

[tree diagram:
IP
DP — Ram
'Big DP' — bought
DP — THAT CD
CorDP — that CD
CP — Which CD is on sale]
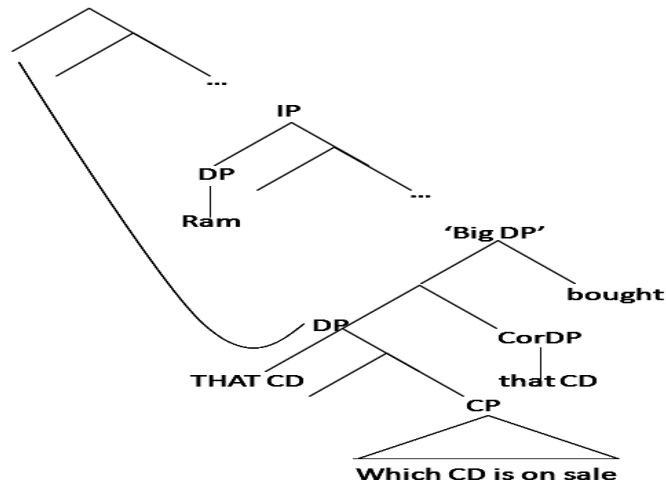
b. 'Which girl which CD heard, that girl that CD bought'

[tree diagram:
CP — Which girl which CD heard
IP
CorDP — that girl
CorDP — that CD — bought]
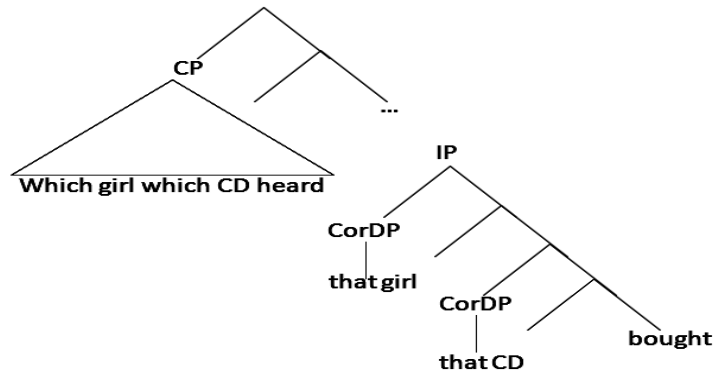
As shown most extensively in Bhatt (2003, 2005), this dual analysis receives support from the fact that in simple, but not in multiple, correlatives the relation between the relative clause and the correlative pronoun is sensitive to islands (Dayal 1996,183; Mahajan 2000, fn.10, and Bhatt 2005); and from the fact that in simple, but not in multiple, correlatives there is obligatory reconstruction of the

fronted relative clause, as evidenced by pronominal binding facts and Principle C violations. For exemplification, see Bhatt (2003, section 3.3.3; 2005).[20]

A further difference between multiple and simple correlatives is represented by the possibility of 'deleting' correlative pronouns when the relative phrases have overt Case. As noted in Bhatt (1997), who attributes the observation to Veneeta Dayal, this is possible in multiple correlatives ((17)) but not in simple correlatives ((18)) (also see Bhatt 2003, section 4):

---

[20] Anderson (2005) makes the interesting observation that Nepali shows a semantic distinction between the two structures (17)a and b. The former is associated with a restrictive (specific) interpretation, the latter with an indefinite (free choice) interpretation. The evidence for this comes from the fact when the correlative is in absolute initial position both interpretations are available while only one, the restrictive (specific) interpretation, is possible when the correlative is adjacent to the correlative pronoun. See (i)a and b:

(i) a. jun manche-lai bhok lag-eko cha, ma us-lai khana din-chu(= Anderson's 2005, ex. (15))
    REL man-DAT hunger attach-PFPT 3SG.PR, 1SG.NOM 3SG.DAT food give-1SG.PR
   either: 'I will give food to the man who is hungry' (specific man – restrictive relative)
     or: 'I will give food to any man who is hungry' (any hungry man – free relative)
 b. ma jun manche-lai bhok lag-eko cha, tyo manche-lai khana din-chu(= Anderson's 2005, ex. (16))
    1SG.NOM REL man-DAT hunger attach-PFPT 3SG.PR, DEM man-DAT food give-1SG.PR
    'I will give food to the man who is hungry' (specific man)

This makes sense, according to Anderson (2005), if the initial position can either be filled by movement of the correlative relative from the internal position adjacent to the correlative DP (which gives the restrictive, specific, interpretation) or by base generating the simple correlative CP (like multiple correlatives) in initial position (which gives the free choice interpretation). It remains to be seen whether this holds of other Indo-Aryan languages as well.

Dayal (1996, chapter 6, section 2) suggests that multiple correlatives in Hindi have a functional reading, which apparently "can also be used to refer to a unique pair of individuals in the contextual domain." (p.204).

Additionally, it should be observed that if simple correlatives can also access the base generated structure of multiple correlatives, they would be expected to show no necessary island sensitivity nor obligatory reconstruction. The facts here are contradictory. While Mahajan (2000,227fn10) and Bhatt (2003, 2005) claim that the correlative pronoun cannot be found within an island (see (ii)), McCawley (2004) gives one case of a correlative pronoun within a relative clause complex NP island judged possible by his informants (his orthography has been uniformized to the one used here). See (iii):

(ii) *[jo si:ta:-ko acha: lagta: hɛ] mɛ [DP yah ba:t [CP ki vo a:dmi: pa:gal hɛ]]
      who Sita-DAT nice seem be-PRES I this fact that that man crazy be-PRES know be-PRES
     'I know the fact that the man who Sita likes is crazy' (= (ii) of fn.10 of Mahajan 2000)

(iii) [jo laRkii vaha khaRii hai], ram ne vo paRha, jo us ne likha
     Which girl there standing is, Ram read the letter that she wrote

Further investigation is needed here, also in relation to the apparent possibility of extracting from correlatives (and *if* clauses) vs. the impossibility of extracting from embedded postnominal and extraposed relatives reported in Dwivedi (1994a,b). Perhaps extraction is possible from the adjunct CP correlative but not from the DP correlative.

Guglielmo Cinque

(17) [jis$_i$ ne   jo$_j$ chahaa]   (us$_i$ ne vo$_j$)   kiyaa
      REL.obl ERG   REL want.Pfv   DEM.obl ERG DEM   do.Pfv
      'Whoever whatever wanted, they did that'   (= (24) of Bhatt 1997,64)

(18) [jis   laRkii=ko Srini pasand hai] *(vo)   khaRii   hai
      REL.obl girl=DAT   S. like be.PRS DEM   standing be.PRS
      'The girl who likes Srini is standing'   (=(9)b of Bhatt 1997,57)

That simple and multiple correlatives should not be treated as a homogeneous construction is also shown by the fact that not all languages having correlatives allow for multiple correlatives. This is the case of Bambara, as reported in Pollard and Sag (1994,229,fn.10) and that of Basque, as reported in Rebuschi (1999,59).

NOTE 4: Non-restrictive correlatives

Dayal (1996), on the basis of the ungrammaticality of examples like (19) below, concludes that Hindi correlatives cannot be non-restrictive "since non-restrictives typically occur with proper names" (p.182).[21]

(19)   *jo laRkii khaRii hai anu lambii hai
        which girl standing be-PR Anu tall is
        'Anu, who is standing, is tall'   (= ex. (43) of Dayal 1996,182)

The question remains whether this is a property of Hindi or of correlatives more generally. To judge from the fact that the closely related Indo-Aryan language Marathi can apparently form non-restrictive correlatives, one has to conclude that the impossibility of (19) in Hindi is not due to some inherent feature of the correlative construction, but is a property of the grammar of Hindi (to be understood). The possibility of non-restrictive correlatives in "rethorical speech and writing" in Marathi is noted in Gupte (1975,77), where such examples as (20)a-b are reported (also see Pandharipande 1997,82f):[22]

---

[21] Also see Gupta (1986,34). The same is claimed by Butt, King and Roth (2007, section 4.2) for the Urdu variant of Hindi/Urdu, and by Bhatia (1993,55) for Punjabi.

[22] The existence of non-restrictive correlatives in Marathi was independently pointed out to me by Avinash Pandey and Renuka Ozarkar. Renuka Ozarkar gave me the following additional example of a non-restrictive correlative in Marathi:

i) ji-ne   maajhyaa-saaThii kaSTa ghet-l-e, tii maajhii aaii aataa jiwanta naahii.
    REL.fem-ERG me-for efforts take-PERF-3P.PLURAL, that my-FEM mother now alive not-PRES
    'My mother, who took efforts for me, is not alive anymore.'

Non-restrictive correlatives were apparently also possible in Sanskrit. See Davison (2009,227).

(20)   a.   jā-nni   gāt^hā racali   te tukārām mahārāj dehulā   janmale
           REL-INSTR Gatha composed that St.Tukaram   Dehu-at   was born
           'St.Tukaram, who composed the Gatha, was born in Dehu'

       b.   gānd^hi-nni   jā-nnā guru   mānale   te gok^hale mawāl   hote
           Gandhi-INSTR REL-to   teacher regarded that Gokhale moderate was
           'Gokhale, whom Gandhi regarded as (his) teacher, was a moderate'

As a matter of fact, given the possibility of resuming a DP followed by a non-restrictive relative clause with a correlative phrase, as in (21) from Bangla, it should in principle be possible, if the language permits it, to 'delete' the external Head like is possible with the external Head of restrictives (cf. (3) and (5) above):

(21)   bhoddrolok, Jini amar ãttio, tini bose achen (Morshed 1986,38)
        Gentleman, who my relative, he sitting is
        'The gentleman, who is my relative, is sitting'

Thus the possibility of non-restrictive correlatives may simply reduce to whether the language allows deletion of the external Head of non-restrictives (Marathi) or not (Hindi).

       Interestingly, non-restrictive correlatives are also attested in other language families. See (22) from Jalonke (of the Central Mande branch of Niger-Congo), and the relative discussion in Lüpke (2005,131f):

(22)   N naaxan a   fala-m'   i bɛ jɛɛ,   n saa-xi saar-ɛɛ ma
       1SG   REL   3SG speak-IPFV 2SG for PART, 1SG lie-PF bed-DEF at
       (lit.) which I is speaking to you now, I lie in bed
       I, who am talking to you now, I am lying on the bed.'

NOTE 5: Correlatives as a non exclusive relativization strategy

To judge from the substantive lists of languages with correlatives given in de Vries (2002,388 and 412), Bhatt (2003,491), and Lipták (2009a,10f) it seems that there may be no single language for which correlatives are the only relativization strategy available. Correlatives invariably appear to co-occur either with embedded postnominal or extraposed relatives (most Indo-Aryan languages, Slavic languages, Warlpiri, etc.), or with prenominal non finite relatives (Dravidian languages, Sinhala, etc.), or with internally Headed relatives (Bambara, Wappo, etc.). From what I have been able to see in the literature on

relative clauses, no language is described as having correlatives as its only type of relative clause.[23]

This fact (assuming it to be a fact) should actually not be surprising if one thinks that simple correlatives (setting multiple correlatives aside, which are no relative clauses) are just left dislocated DPs containing a relative clause of one or another of the existing types (externally Headed postnominal, externally Headed prenominal, internally Headed, and Headless or free) resumed by a phrase in the main clause.

References

Almeida, Matthew, S.J. 1989. *A Description of Konkani.* Panaji: Thomas
    Stephens Konknnikendr.
Ananda, Lalith M.G. 2008. *The Cleft Construction in Sinhala*. MPhil.
    Dissertation. Jawaharlal Nehru University, New Delhi.
Anderson, Corinna. 2005. Two Types of Nepali Correlatives. In Y.Yadava,
    G.Bhattarai, R.R. Lohani, B.Prasain and K.Parajuli, eds., *Contemporary Issues in
    Nepalese Linguistics*. 1-12. Kathmandu: Linguistic Society of Nepal.
Andrews, Avery D. 1975. *Studies in the Syntax of Relative and Comparative
    Clauses*. Ph.D.thesis, Massachusetts Institute of Technology.
Bagchi, Tista**.** 1994. Bangla correlative pronouns, relative clause order and D‐linking.
    *In M. Butt et al. (eds) Theoretical Perspectives on Word Order in South Asian
    Languages* (CSLI Lecture Notes 50) 13–30. Stanford, CA: CSLI.
Bertone, Carmela. 2006. La struttura del sintagma determinante nella Lingua dei Segni
    Italiana (LIS). Ph.D. Dissertation, University of Venice.
Bhatia, Tej K. 1993. *Punjabi*. London: Routledge.
Bhatt, Rajesh. 1997. Matching effects and the syntax–morphology interface: evidence from
    Hindi correlatives. In B.Bruening (ed.) *Proceedings of SCIL 8* (MIT Working
    Papers in Linguistics 31) 53–68. Cambridge, MA: MITWPL.
Bhatt, Rajesh. 2003. Locality in Correlatives. *Natural Language and
    Linguistic Theory* 21.485-541.
Bhatt, Rajesh. 2005. Correlative Clauses. Handout of a course at the 2005 LOT Summer
    School, Leiden (June 16) (http://people.umass.edu/bhatt/752-s05/n44.pdf)
Bhatt, Rakesh Mohan. 1999. *Verb Movement and the Syntax of Kashmiri*. Dordrecht:
    Kluwer.
Bianchi, Valentina. 1999. *Consequences of Antisymmetry: Headed Relative Clauses*. Berlin:
    Mouton de Gruyter.

---

[23] Actually, Creissels (2009,43) states that "[l]e malinké n'a pas de relatives adnominales: les seules relatives du malinké sont les relatives correlatives [..].", but, as he makes clear, the correlatives of Malinké are left dislocated Internally Headed relatives, which in contrast to the closely related language Bambara (cf. fn.19 above), appear not to be able to occur in argument position (Creissels 2009,51). This, if true, remains to be understood.

Bird, Charles B. 1968. Relative Clauses in Bambara. *The Journal of West African Languages* 5.35-47.

Branchini, Chiara and Caterina Donati. 2009. Relatively different. Italian Sign Language relative clauses in a typological perspective. In A. Lipták, ed., *Correlatives Cross-Linguistically*. 157-191. Amsterdam: Benjamins.

Brunelli, Michele (2006) *The grammar of Italian Sign Language, with a study about its restrictive relative clauses*. MA Thesis, University of Venice (http://michelebrunelli.interfree.it/LISgrammRestrRel.html).

Butt, Miriam, Tracy Holloway King, and Sebastian Roth. 2007. Urdu Correlatives: Theoretical and Implementational Issues. In M.Butt and T.Holloway King (eds.) *Proceedings of the LFG07 Conference*. Stanford: CSLI Publications (http://csli-publications.stanford.edu/LFG/ 12/lfg07.pdf).

Cable, Seth. 2007. The Syntax of the Tibetan Correlative. Manuscript. University of British Columbia. Vancouver, BC. (http://people.umass.edu/scable/papers/Tibetan-Correlative-Syntax.pdf). (appeared in A.Liptak 2009b).

Carlson, Greg N. 1977. Amount Relatives. *Language* 53, 520–542.

Cecchetto, Carlo, Carlo Geraci and Sandro Zucchi. 2006. Strategies of relativization in Italian Sign Language. *Natural Language and Linguistic Theory* 24.945-975.

Cinque, Guglielmo. 1990. *Types of A-Bar Dependencies*. Cambridge, Mass.: MIT Press.

Cinque, Guglielmo. 2008. More on the indefinite character of the Head of restrictive relatives. In P.Benincà, F.Damonte and N.Penello, eds., *Selected Proceedings of the 34th Incontro di Grammatica Generativa*, Unipress, Padova (Special issue of the *Rivista di grammatica generativa*, vol. 33, 3-24 (http://dspace-unive.cilea.it/handle/10278/880).

Cinque, Guglielmo. 2009. The Fundamental Left-Right Asymmetry of Natural Languages. In S.Scalise, E.Magni, A.Bisetto, eds., *Universals of Language Today*. Dordrecht: Springer (2009), pp. 165-184 (http://dspace-unive.cilea. it/handle/10278/214).

Cinque, Guglielmo. In preparation. *A unified analysis of relative clauses*. University of Venice.

Creissels, Denis. 2009. Les relatives corrélatives: le cas du malinké de Kita. Langages 174.39-52

Dasgupta, Probal. 1980. Questions and relative and complement clauses in a Bangla grammar. Ph.D. Dissertation. New York University.

Dasgupta, Probal.2006. Unifying relativization and control in Bangla. In M. Banerjee et al., eds., *Nyaya-Vasistha: Felicitation Volume of Professor V.N. Jha*. 138-170. Kolkata: Sanskrit Pustak Bhandar.

Davison, Alice**.** 2009. Adjunction, features and locality in Sanskrit and Hindi/Urdu correlatives. In A. Lipták, ed., *Correlatives Cross-Linguistically*. 223-262. Amsterdam: Benjamins.

Dayal, Veneeta. 1996. *Locality in wh-Quantification*. Dordrecht: Kluwer.

Downing, Bruce T. 1973. Correlative relative clauses in universal grammar. *Minnesota Working Papers in Linguistics and Philosophy of Language* 2.1-17.

Downing, Bruce T. 1978. Some Universals of Relative Clause Structure. In J.Greenberg, ed., *Universals of Human Language*. *Vol.4*. 375-418. Stanford, CA: Stanford University Press.

Dwivedi, Veena. 1994a. Topicalization in Hindi and the Correlative Construction. In M.Butt, T.Holloway King and G.Ramchand (eds.) *Theoretical Perspectives on Word Order in South Asian Languages*. 91-118. Stanford: CSLI Publications.

Dwivedi, Veena. 1994b. *Syntactic Dependencies and Relative Phrases in Hindi*. Ph.D. Dissertation, University of Massachusetts, Amherst.

Gorbet, Larry. 1977. Headless relative clauses in the southwest: are they related? *BLS* 3.270-78

Grosu, Alexander. 2002. Strange relatives at the interface of two millennia. *Glot International* 6.6.145-167.

Gupta, Sagarmal. 1986. *Discourse Grammar of Hindi. A Study in Relative Clauses*. New Delhi: Bahri Publications Private Limited.

Gupte, Sharad M. 1975. *Relative Constructions in Marathi*. Ph.D. Dissertation, Michigan State University.

Hale, Kenneth. 1976. The adjoined relative clause in Australia. In R. M. W. Dixon, ed., *Grammatical Categories in Australian Languages*. 78-105. Canberra: AIAS

Hook, Peter Edwin and Omkar Nath Koul. 1996. Kashmiri as a V-2 language. In V.Swarajya Lakshmi and A.Mukherjee, eds., *Word Order in Indian Languages*. 95-105. Hyderabad: Center for Advanced Study in Linguistics, Osmania University and Booklinks Corporation.

Izvorski, Roumyana (Pancheva). 1996. The syntax and semantics of correlative proforms. *NELS* 26.133-147

Izvorski, Roumyana (Pancheva). 2000. Free Adjunct Free Relatives. *Proceedings of WCCFL 19*.232-245.

Junghare, Indira. 1973. Restrictive Relative Clauses in Marathi. *Indian Linguistics* 34.4.251-262.

Kayne, Richard S. 1972. Subject inversion in French interrogatives. In J.Casagrande and B.Saciuk, eds., *Studies in Romance Languages*. 70-126. Newbury House, Rowley, Mass.

Kayne, Richard S. 1994. *The Antisymmetry of Syntax*. Cambridge, Mass.: MIT Press.

Kayne, Richard S. 2004. Here and There. In C.Leclère, É.Laporte, M.Piot, M.Silberzstein, eds, *Lexique, Syntaxe et Lexique-grammaire. Syntax, Lexis & Lexicon-Grammar. Papers in Honour of Maurice Gross*, 275-285. Amsterdam: Benjamins (also in Kayne 2005b, pp.65-84).

Kayne, Richard S. 2005a. Some Notes on Comparative Syntax, with Special Reference to English and French. In G.Cinque and R.S.Kayne, eds., *The Oxford Handbook of Comparative Syntax*, 3-69. New York: Oxford University Press (also in Kayne 2005b, pp. 277-333).

Kayne, Richard S. 2005b. *Movement and Silence*. New York: Oxford University Press.

Kayne, Richard S. 2007. On the Syntax of Quantity in English. In J.Bayer, T.Bhattacharya and M.T. Hany Babu, eds., *Linguistic Theory and South Asian Languages. Essays in*

honour of K.A.Jayaseelan. 73-105. Amsterdam: Benjamins (also in Kayne 2005b, pp.176-214).

Keenan, Edward. 1985. Relative Clauses. In T.Shopen, ed., *Language Typology and Syntactic Description.* Vol.II. *Complex Constructions*. 141-170. Cambridge: Cambridge University Press.

Koul, Omkar N. 2003. Kashmiri. In G.Cardona and D.Jain, eds., *The Indo-Aryan Languages*. 895-952. London: Routledge.

Krapova, Iliyana. 2009. Bulgarian Relative and Factive Clauses with an Invariant Complementizer. Ms., University of Venice (To appear in *Lingua*.)

Lakshmi Bai, B. 1985. Some notes on correlative constructions in Dravidian. In: *For Gordon H. Fairbanks*, ed. by V. Z. Acson & R. L. Leed, 181-90. Honolulu: University of Hawaii Press.

Leung, Tommi T.-C. (2007a) *Syntactic Derivation and the Theory of Matching Contextual Features*. Ph.D. Dissertation, University of Southern California.

Leung, Tommi T.-C. (2007b) "Correlatives and the conditions on chain formation" *Proceedings of the International Conference on Linguistics in Korea* (ICLK 2007), Seoul National University, Seoul, KGGC.
(http://tommi.leung.googlepages.com/ICLK2007-handout.pdf).

Leung, Tommi T.-C. 2007c. On the Typology of Correlative Constructions. Handout of a paper presented at the conference of the Association of Linguistic Typology (ALT VII) CNRS, Paris, France (http://tommi.leung.googlepages.com/ALTVIIhandout.pdf).

Lipták, Anikó. 2004. On the Correlative Nature of Hungarian Left-peripheral Relatives. In B. Shaer, W. Frey, C. Maienborn (eds), *Proceedings of the Dislocated Elements Workshop (ZAS Berlin; November 2003), ZAS Papers in Linguistics* 35. 1: 287-313. Berlin: ZAS.

Lipták, Anikó, ed., 2009a. The landscape of correlatives: An empirical and analytical survey. In A. Lipták, ed., 2009b. Pp.1-46.

Lipták, Anikó, ed., 2009b. *Correlatives Cross-Linguistically*. Amsterdam: Benjamins.

Lüpke, Friederike. 2005. *A Grammar of Jalonke Argument Structure*. Ph.D. Dissertation, Radboud Universiteit, Nijmegen (Vol.30 of the Max Planck Institute in Psycholinguistics Series) (http://repository.ubn.ru.nl/handle/2066/19598 or https://eprints.soas.ac.uk/4853/1/gramofjaa.pdf).

Mahajan, Anoop. 2000. Relative Asymmetries and Hindi Correlatives. In A.Alexiadou, P.Law, A.Meinunger, C.Wilder (eds.) *The Syntax of Relative Clauses*. 201-229. Amsterdam: Benjamins

McCawley, James D. 2004. Remarks on Adsentential, Adnominal and Extraposed Relative Clauses in Hindi. In V.Dayal and A.Mahajan, eds., *Clause Structure in South Asian Languages*. 291-311. Dordrecht: Kluwer.

Milsark, Gary. 1974. *Existential sentences in English*. Ph.D. Dissertation, MIT.

Morshed, Abul K.M. 1986. *Relativization in Bengali*. Dhaka (Bangladesh): University of Dhaka (originally a dissertation from the University of Edinburgh).

Neukom Lukas and Manideepa Patnaik. 2003. *A grammar of Oriya* (ASAS, Nr.17). Zürich: Seminar für Allgemeine Sprachwissenschaft, Universität Zürich.

Pandharipande, RajeshwariV. 1997. *Marathi*. London: Routledge.

Guglielmo Cinque

Pollard, Carl and Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar*. Chicago: The University of Chicago Press.

Postal,Paul. 1971. *Crossover Phenomena.* New York: Holt, Rinehart and Winston.

Rebuschi, Georges. 1999. 'Types de langues et types de constructions: le cas des corrélatives' ; *LINX*, numéro spécial thématique *Typologie des langues, universaux linguistiques*, pp.55-72.

Rebuschi, Georges. 2009. Position du basque dans la typologie des relatives corrélatives. *Langages* 174.25-38.

Shukla, Shaligram. 1981. *Bhojpuri Grammar*. Washington, D.C. : Georgetown University Press.

Singh, Udaya Narayana. 1980. Relative Clause Formation in Maithili. *Nepalese Linguistics* 1.27-39 (http://www.ciil.org/Main/Faculty/uday.html).

Srivastav, Veneeta (Dayal). 1988. Relative Clauses in Hindi and Learnability. In V.Srivastav, J.W.Gair, and K.Wali (eds.) *Papers on South Asian Linguistics* (Special Issue of *Cornell University Working Papers in Linguistics* 8.133-160).

Srivastav, Veneeta (Dayal). 1991. The syntax and semantics of correlatives. *Natural Language and Linguistic Theory* 9.637-686.

Thompson, Sandra A., Park, Joseph Sung-Yul, and Charles N. Li. 2006. *A Reference Grammar of Wappo.* Berkeley: University of California Press (http://repositories.cdlib.org/ucpress/ucpl/vol_138/).

Tiffou, Étienne and Richard Patry. 1995. La relative en bourouchaski du Yasin. *Bulletin de la Société de Linguistique de Paris,* 90.335-91.

Uriagereka, Juan. 1995. Aspects of the Syntax of Clitic Placement in Western Romance. *Linguistic Inquiry* 26.79-123.

Vries, Mark de. 2002. *The Syntax of Relativization*. Ph.D. Dissertation, University of Amsterdam.

Wali, Kashi. 1982. Marathi Correlatives: A Conspectus. In P. J. Mistry (ed.), *South Asian Review: Studies in South Asian Languages and Linguistics*. Jacksonville, Florida, South Asian Literary Association. 78-88. (reprinted in K.Wali. 2006, pp. 288-297).

Wali, Kashi. 2006. *Marathi*. Delhi: Indian Institute of Language Studies.

Zribi-Hertz, Anne and Jean-François Hanne. 1995. Pronoms, determinants et relatives en bambara de Bamako. *Linguistique africaine* 15.91-136.

Discourse and Generative Grammar: a substantivist approach

Probal Dasgupta
Indian Statistical Institute, Kolkata

Abstract

The fresh assembly property of individual sentences does not logically follow from the infinity of the set of sentences. The phraseological fact that certain subsentential constituents are precompiled invites a multiple validation approach to the phenomenon of authorized fresh assembly. Multiple validation – the key manoeuvre of substantivism – enables principled treatment of arbitrariness and opacity issues in Word Formation Strategies (a familiar device that handles pure morphology), Word Extension Strategies (which handle clitics) and Phrase Formation Rules (which introduce compositional but non-generalizable phrases).

1. Introduction

In this paper, section 1 develops some proposals in the context of generative grammar and the study of discourse. On the basis of this set of 'substantivist' proposals, section 2 considers issues of arbitrariness and opacity in relation to concrete examples and highlights the role of multiple validation in substantivist inquiry. Multiple validation issues connect section 3, on strategies and strategy shadow, with section 4, on the proper treatment of constructions that elude the net of the regular syntax.

One route that leads from garden variety generative grammar to questions of discourse is a redundancy argument. Consider passages (1) and (2):

(1) John came in. He coughed. He apologized.
(2) John came in, coughed, and apologized.

Assuming that *he* refers to John, we can regard passages (1) and (2) as statably equivalent. A formal statement of this equivalence must take on board what there is to say about the three-sentence discourse (1), and yet rigorously correspond to the relevant formal account of the single sentence (2). Thus, certain exact statements about discourse have to do some formal work that would ordinarily fall under syntactic description. Those formal statements – perhaps in an Exact Discourse Analysis module EDA (or in a Discourse-Syntax Interface module DSI) – should take charge of much of the formal description of clausal co-

Probal Dasgupta

/subordination. The syntax of sentences should thus leave it to that module to handle not just three-sentence discourses like (2), but also a sentence like (1) that coordinates three clauses.

This argument may lead some of us to conclude that whatever falls within the purview of EDA must be removed from sentence syntax. So construed, this reasoning becomes a redundancy argument. Indeed, the formalistic approach to exact linguistic description does employ redundancy arguments. For it favours the unique allocation of particular characterization tasks to one module or to another. But this is not the only way to do formal linguistics.

When we take up the formal study of language, we have to choose between two different conceptualizations of the task of characterizing linguistic phenomena in terms of several modules of the grammar. The **formalist** conceptualization of this task in formal linguistics favours the unique allocation of particular characterization tasks to one module or to another. For example, if one proposes to derive the English comparative adjective *taller* from a syntactic representation isomorphic to the phrasal comparative *more complex*, one is expressing a formalist desire to make 'the' statement about English comparatives only once in the grammar.

The **substantivist** conceptualization of the task of formal characterization of linguistic phenomena (Dasgupta 1977, 1989, 1993, 2005, 2006, Dasgupta, Ford and Singh 2000) permits and often encourages co-characterization or multiple characterization. Substantivism works with the intersecting economies of distinct modules to find, by triangulation, what principles of economy are actually operative in particular phenomena one is describing. Unsurprisingly, the first explicit articulation of the substantivist programme (Dasgupta 1989) emerged in a translation studies context. Substantivist inquiry mediates, translates, between formal apparatuses, between modules, between languages, to co-characterize particular phenomena in keeping with the principles of economy governing them. For English comparatives substantivist principles favour a morphological description of *taller* and a syntactic description of *more complex*, with a shared grammatical feature [+comparative] connecting the two.

In all frameworks, linguists agree that descriptions have to implement a preference for specific forms over general forms, for instance by arranging for the more specific form *taller* to preempt or 'block' the more general form *more tall* (if we set aside cases like *more tall than wide*). There are certain unresolved questions, however, about what does or doesn't get preempted, and why. Hence this paper. To see how the issue plays out, consider (3) and (4):

(3) What say you?

(4) What do you say?

If we had an extensive irregular pattern on our hands – such as *What think you? What want you? What like you?*, alongside the regular *What do you think?* etc. – we would subsume the (3)-(4) pair under a proper story. But (3) is isolated; so is (5); and their presence in English does not block the regular questions (4) or (6):

(5) What have we here?
(6) What do we have here?

How does the intersecting economies idea manage the coexistence of isolated irregular forms with the rest of the system?

The syntax proper, however a particular framework may choose to run its version of the *Do*-Support routine, licenses (4) and (6) as the routine forms. Elsewhere, a clump of special idiosyncratic expressions built around the verb *to say* – such as *Says who?, What say you?,* the schema *Said like a(n) X* (X = *physicist, good Christian, engineer*, or some other stereotype), plus another expression or two – live in the suburbs of the syntax. They do not have the grossly irregular look of rote-learnt opaque expressions like *Easy come, easy go* or *every which way*. Our *say*-expressions are easy to parse, and *Said like a(n) X* is not even grammatically deviant, but merely elliptical. At the same time, what authorizes them is the special dispensation for the clump of *say*-expressions. The syntax does not license these formulaic expressions. It only issues a special suburban pass for them, so to speak.

Moving from this informal metaphor of issuing passes to serious formalization involves breaking this theme down into operational questions. The licensing of semi-regular or irregular entities is a problem that the syntax and the morphology both have to deal with. To regard discourse as a neighbour only of the syntax module is to seriously misconstrue the issues. On the syntagmatic axis, it makes some sense to say that discourse begins where syntax ends. But on the paradigmatic plane, discursive relations obtain between any potential utterance – any word/ phrase/ sentence – and its significant others, i.e. the words/ phrases/ sentences that systematically resemble and/or differ from the item in question. Given these considerations, when we wish to understand better how language manages the coexistence of isolated irregular forms with regular forms, it becomes perfectly natural to move to the morphology in search of illumination.

Our discussion of morphology assumes some familiarity with the approach known as Whole Word Morphology, WWM (see the standard exposition in Ford,

Singh & Martohardjono 1997). We return shortly to the task of contextualizing its basic assumptions within the substantivist approach; see our exposition from (10) onwards. Under WWM assumptions, it is the Word Formation Strategies (WFSs) of English morphology that license regular plurals: *foxes, girls, fans, persons, brothers*. The lexicon extrastrategically authorizes irregular plurals like *oxen, children, men, people, brethren*. Our metaphoric statement that the morphology issues suburban passes to these translates into a proposal that the morphology should bring these items under a rigorous regime of secondary licensing by applying a wild card type Word Formation Strategy. Accordingly, we hereby propose this addition to the WWM toolkit:

(7) Secondary Licensing (morphology)

$[X]_\alpha \longleftrightarrow [X^w]_\beta$, where:

X is a word,
X^w is a word that wildly differs from X formally, and
α and β are feature bundles that figure in regular WFSs in the language (see (10) below)

The empirical content of our proposal is that, since English has no regular optatives or desideratives or duals, it will not be allowed to have any irregular optatives or desideratives or duals. *Was, were*, and *am* may strike some readers as counterexamples. But number agreement is available in English morphology, as is agreement marking for the first person singular (*I go* contrasts with *he goes*).

In the syntax, the equivalent move is for a lexico-phrasally authorized formation to receive rigorously Secondary Licensing in the same wild card mode, as shown at (8). Proposal (8) leads to empirical prediction (9):

(8) Secondary Licensing (syntax)

Assign wild card features to the illocution, drawing from the UG set of features available to illocutions

(9)   (a) Secondary Licensing in the syntax will target root sentences

(b) Formations so licensed will not push regular competitors into ill-formedness

Why does (9) follow from (8)? Because, (a), the clause type information of an ordinary, non-'root'-type embedded clause is so completely specified by matrix forces that wild card respecification gets preempted. And (b), Secondary Licensing routes the irregular sentence straight from lexico-phraseological

storage into sentential use. Bypassing the syntax machinery means that that system continues to generate its routine products without disturbance or upstaging.

How does the syntactic assembly cooperate with lexical storage to give such items a semi-transparent look, though? Careful answers to such a question will have to be framework-specific. On 'left periphery' assumptions, for instance, one may underspecify nodes in the C[omplementizer] region of the clausal architecture, assigning an interrogative feature to one of them, perhaps to Force[0]. UG principles guarantee that an interrogative in the body of the clause (the *what* in *What say you?*, for example) featurally interacts with the C system, but that the auxiliary-fed special effects characteristic of English are absent because of underspecification in the C system. Translating this account into other frameworks is a straightforward matter. Issues of transparency and opacity reappear on our screen in section 2.

We can now situate the WWM framework of morphology (Ford, Singh and Martohardjono 1997) in the substantivist research programme.

One fundamental manoeuvre that distinguishes substantivist from formalist inquiry is the substance-focused use of multiple characterization. Here the term substance refers – along the lines of Chomsky & Halle (1968: chapter 9) – to factors that determine patterns of marked vs unmarked, natural vs unnatural, easy vs difficult, basic vs non-basic, basilectal vs acrolectal, and even spoken vs written. The leading question in the substantivist study of language connects substance with multiple characterization as follows. Lexical storage maximizes ease of retention and access; syntactic processes maximize ease of assembly; there are also other maximizations at work. How do linguistic representations, in explicit compliance with imperatives emanating from several modules, manage these intersecting economies in such a way that a base language – easy for the child to acquire – can exist? What extensions associate this base with the full richness of human language?

The bit about explicit compliance obliges each linguistic string to wear multiple representations accountable to multiple sources of validation. All doctrines agree that every string must dress up for sound and for meaning and must be okayed in both of those dimensions. Substantivism says, in addition, that a string must invite and receive validation also from Mod_1, Mod_2, …, Mod_n – the morphology, the syntax, and any other relevant modules. This connects with the markedness legacy at the tenet that the business of managing a viable base involves, for every i, showing the string's full Mod_i-face to the specific principles

of the module Mod$_i$. For example, even in a language where the syntax guarantees that to the left of every noun there is a determiner, nonetheless semantic representations must fully specify every determiner instead of leaving blanks to be filled in by the syntax.

Formalist grammarians, who despite a declaration or two seem not to have taken markedness on board, have consistently found it convenient to apply pre-markedness 'formal economy' criteria that seek mechanical generalities and that make it look reasonable to postulate 'roots' and 'affixes' in their morphology. They accordingly continue the structuralist habit of subsuming these under a superordinate category of 'morphemes', thereby proposing that roots and affixes have serious properties in common. Even 'distributed morphology' solutions of formalist devising, while their precompilations sometimes resemble the substantivist's integral words at a superficial level, invoke 'vocabulary items' that keep the spirit of morphemics alive. Thus, all versions of formalist linguistics leave open the possibility that languages may exist which suspend or reverse the natural asymmetry between content-denoting lexical materials and function-signalling modifications of these.

Before we plunge into the formal work of this section, we pause to unpack this crucial empirical point about asymmetry. To see what is at stake, imagine a world in which some languages suspend or reverse the asymmetry between the formalist's 'roots' and 'affixes', i.e. between lexical material and its function-signalling modifications. The exercise may run as follows.

In the real world, a believer in 'morphemes' finds that in English the 'past tense affix' is manifested in *sighed* as the segment /d/ but in *blew, threw* as a 'replacive' – as a substitution of /u:/ for /ou/. Suppose we turn this around; and note that the exercise remains materially unaltered if we move from morphemics to precompilation accounts of the 'distributed' kind. Imagine a world, then, in which what we shall call the *Spenglish* past tense 'affix' has the shape /éd/ nearly everywhere and carries primary stress. This involves imagining that Spenglish words are mostly spellED as in our English but spelling-pronouncED with primary stress on /éd/. Now we get to the hard part of this exercise: let there be a quirky Spenglish verb /tu éi/ 'to eat', and let this verbal 'root' morpheme be manifested as a replacive when it interacts with the 'affix morphemes' /éd/ and /íŋ/. Thus we get /eid/ for 'ate', where the 'root morpheme' appears as a 'replacive allomorph' that substitutes /ei/ for /e/. Likewise, in order to express 'eating', a Spenglish speaker would say /i:ŋ/, where the 'root morpheme' again shows up as a 'replacive', as the substitution of /i:/ for /i/.

To flesh the picture out, we add another Spenglish verb /tu éi/. It means 'to drink' and interacts opaquely with the 'affixes' to yield /eəd/ 'drank' and /iəŋ/

'drinking'. Thus, an /e/ $\rightarrow$ /eə/ 'replacive allomorph' manifests this 'root morpheme' in the past form; there is an /i/ $\rightarrow$ /iə/ manifestation in the gerund form.

The reader knows, of course, that languages in fact never give lexical material so little space and function-signalling modifications so much. But surely some languages could have done so, if 'roots' and 'affixes' had indeed been equal, and thus validly subsumable under a superordinate 'morpheme' notion. How are we to respond to the emphatic absence, in the real world, of such interchangeability between 'roots' and 'affixes'?

Substantivist inquiry's core commitment keeps it focused on optimality or 'markedness' factors that determine fundamental asymmetries in linguistic phenomena. This focus makes it important to formulate the basic concepts of morphology in a way that makes "Can 'affixes' in language B behave the way 'roots' do in language A?" an unformulable question – for instance, by prohibiting reference to 'roots', 'affixes', 'morphemes', or euphemisms like 'vocabulary items' (if by this one means anything smaller than words). One framework with this property, WWM, is built around essentially the following conceptualization of what constitutes a Word Formation Strategy:

(10)　(a) $[X]_\alpha \leftrightarrow [X']_\beta$

(b) Schema (a) states that at least two pairs $(X_1, X'_1; X_2, X'_2)$ of words in the mental lexicon of speaker S anchor a correspondence that has the properties specified here;

(c) /X/ and /X'/ are words, the prime and the arrow indicate a bidirectional X-X' mapping, and the form of each side as well as the mapping is specified with appropriate maximization of specificity and generality;

(d) in particular, the representations of /X/ and of /X'/ specify only those phonic features that automatic phonology cannot predict;

(e) α and β are bundles of grammatical features;

(f) formal correspondences as in (a) are associated with interpretation mappings.

Scholars interested in developing some other morphological framework so that it meets the 'root'-'affix' asymmetry challenge at least as seriously as (10) does will no doubt formulate such proposals. WWM assumptions predict the impossibility of a Spenglish by making the 'generalizations' governing /oug, eid,

i:ŋ, ei, eəd, iəŋ/ unformulable. In the absence of other viable proposals, it is Word Formation Strategies in the sense of (10) that substantivist inquiry will stay focused on.

The WWM literature standardly adds to (10a-f) the further assertion that, if X = X′, then α ≠ β. Formulation (10) omits that line and makes lexical correspondence a reflexive relation. Preventing the vacuous or trivial use of mechanisms is a topic to which we shall return in section 3. Right now our task is to begin to situate WWM in the substantivist programme. What has all this to do with discourse, for instance?

Substantivism's debate with formalism turns on the notions of language as code and language as discourse. Even radical forms of formalist linguistics – with multiple spell-out from anarchically plural work-spaces, warps, and other apparently open processes – seek closure, at the level of the finally assembled output of such processes, in the structuralist notion of a composite sign consisting of structured simple signs. For a formalist even a multiclausal sentence is a huge composite sign; only above the sentence level does discourse begin.

It is in this sense that the formalist research programme in linguistics is committed to viewing language as a code, an array of signs, under the assumption of signifier-signified colligation at the sign, and under the generative extension of this view to a non-trivially infinite array of sentence-length composite signs. Lexicalist/ representationalist alternatives to the transformational/ derivational mainstream – including such work on lexical integrity as Aronoff 1976, 1993 – cleave to the language-as-code assumption even when they question the morpheme, or word-internal derivations, or other proposals that have been made within particular implementations but do not belong to formalism's core agenda.

Substantivism keys language at all levels into discourse while continuing to use formal mechanisms. In substantivist inquiry, discourse is the domain of encounter between potential speaker A and potential speaker B, of contact between possible speaking P and possible speaking Q. This makes even word to word relations discursive, for an utterance can be as short as a word. WWM per se offers a formal account of word relatedness patterns. Substantivism chooses the WWM account in part for reasons of its own (such as the argument from Spenglish), and uses this account to, among other things, conceptualize the access that a speaker using word P continuously has to paradigmatic neighbours Q, R, S – in a space whose situatedness in discourse is an independently important object of cognitive science inquiry. Obviously the substantivist take on paradigmatic relations in such discursive space also looks at phrases and clauses, as will become clearer in the context of empirical material taken up in section 3.

This is why the formal workings of (7) in the morphology and of (8) in the syntax raise the stakes for all linguists.

In the morphology, irregular plurals or pasts sometimes preempt regular forms (*men* in English upstages *mans*, and *went* blocks *goed*) and sometimes do not (*people* coexists with *persons*, as does *dreamt* with *dreamed*). In the syntax, however, *Says who?* and *What say you?* do not block the regular *Who says?* and *What do you say?*. We shall return in section 2 to some reasons for complicating this empirical picture. For the moment, though, we take it that this contrast simply polarizes the morphology and the syntax. To the extent that it does, what sense may we make of the phenomena in terms of intersecting economies?

We have already – at the discussion at (8) above – answered the question of why secondary licensing in the syntax does not preempt the primarily licensed regular form. We turn now to the morphological question.

Cases like *persons* vs *people,* or *brothers* vs *brethren,* resemble *elder* vs *older.* Semantically differentiated doublets provide a niche for special forms. *People* and *brethren* exemplify the pluralia tantum phenomenon. *Elder* is similar. You can say *The third and the fifth persons/ brothers,* but not *The third and the fifth people/ brethren*. You can say *older than,* but not *elder than*.

Clear examples like *went* upstaging **goed*, or *gave* blocking **gived*, are to be contrasted with *dreamt* and *learnt* freely alternating, for some speakers, with *dreamed* and *learned*. Such clear cases of irregularity, devoid of semantic doublet properties, call for comment.

One generalization that suggests itself runs as follows. Whenever a lexically isolated irregular form receives secondary licensing in the morphology, it always blocks the regular template: *men, women, children, oxen, hurt (*past*), went* block the expected regular forms *mans, womans, childs* etc. When we find a WFS in competition with a more general WFS, there is sometimes free variation (*dreamt, leant, leapt ~ dreamed, leaned, leaped; learnt, burnt ~ learned, burned*) and sometimes blocking (*meant, crept, slept*). In other words, a free variation pattern implies two competing WFSs, while morphological secondary licensing of a lexical loner implies blocking.

This first approximation does not quite work, though. In Bangla verb morphology (Dasgupta 2001: 166, 171), certain free alternations such as *guchono ~ gochano* 'to arrange neatly', *bulono ~ bolano* 'to stroke', *upRono ~ opRano* 'to uproot', *Sudhrono ~ Sodhrano* 'to correct' are indeed associated with a WFS. But the free alternation between *pechono* and *pichono* 'to step back' is lexically isolated – we find no WFS competition here. The claim we need to make requires

Probal Dasgupta

a more careful formulation – "an *irregular* form bearing a syntactically significant feature freely alternates with a *regular* form only if a less general WFS supports the irregular alternant". The Bangla data just mentioned do not counterexemplify **this** claim, for neither of the alternants in the case of *guchono ~ gochano* 'to arrange neatly' is more regular than the other.

Where we can check this claim in Bangla – a future tense paradigm given in (11), where WFS (11g) supports the irregular variants (11c, f) – the claim is confirmed, whereas the lexically isolated past at (12f) blocks the regular template, a correlation our conjecture leads us to expect:

(11) (a) de 'give!'
  (b) debo '(I) will give' (regular)
  (c) dobo '(I) will give' (irregular)
  (d) ne 'take!'
  (e) nebo '(I) will take' (regular)
  (f) nobo '(I) will take' (irregular)
  (g) $[Ce]_{V,Imp}$ ←→ $[Cobo]_{V,Fut,1p}$

(12) (a) kha 'eat!'
  (b) khelo '(s/he) ate' (regular)
  (c) pa 'get!'
  (d) pelo '(s/he) got' (regular)
  (e) ja 'go!'
  (f) gElo '(s/he) went' (irregular)
  (g) *jelo (supplanted regular form)

If this generalization is able to handle the gross patterns considered so far, we are ready to deal with the more intricate facts that a closer look brings into view – in section 2.

2. Arbitrariness

So far, we have said only that syntax assembles sentences. That such assembly counts as primary licensing of the products so assembled is implicit in the statement that secondary licensing in the sense of (8) takes precompiled material like *What say you?* straight from lexico-phrasal storage into actual use as a sentence – bypassing the regular syntactic assembly process, avoiding competition with its mechanisms, and ensuring that its normal outcome *What do you say?* also counts as well-formed. However, even a regular sentence cannot be compositionally assembled all the way down to its phonological segments. What preexisting items does the syntactic process assemble when it composes a

sentence? What non-transparent input serves as the point of departure for this transparent process of assembly?

If the answer is that in *John loves Mary* the non-transparent input items are essentially the words *John, loves,* and *Mary*, notice that linguists providing this answer do not speak of "opacity" – the natural antonym for "transparency" – to describe these items. One says instead that *John loves Mary* differs from *Jean aime Marie* because the linguistic sign is "arbitrary" by nature.

There is more to this than meets the eye. The concept pair "arbitrary/ motivated" has much in common with the "opaque/ transparent" pair. But it pays to also take a careful look at some differences – one of our tasks in this section.

Formalism's code approach is closely associated with the pair "arbitrary vs motivated". In this sense, it inherits structuralism's tendency to stress relations between signs within a formally structured whole. While substantivism's discourse approach does bring the concepts "opaque vs transparent" to the fore, our proposal is not that these should supersede "arbitrary vs motivated", but that there should be a division of explanatory labour. We wish to use "arbitrary/ motivated" to manage lexical storage and "opaque/ transparent" to drive the fresh assembly of utterances. Clarifying this division of labour, and the proposal that the two economies intersect, will involve a little bit of theorizing. But we are happy to announce that the moves made here will lead to empirical consequences before the section is done.

Recall, from the textbooks linguists cut their teeth on, that simple signs such as French *chien* 'dog' or Bangla *kukur* 'dog' are termed "arbitrary" because no biological or other foundation underwrites their concrete forms. It is possible to slide from the "they are ungrounded" version to the "they constitute the ground" version of this doctrine and thus to view *kukur* as unmotivated – as carrying no clues – while the "relatively motivated" *kukurer* 'dog's' invites comparison with *kukur* 'dog', *beRaler* 'cat's' and *beRal* 'cat'. In such a perspective, arbitrariness and motivation count as natural opposites. One is saying that *kukurer* 'dog's' is relatively motivated because it shares something with *kukur* 'dog' and something else with *beRaler* 'cat's', whereas *kukur* 'dog' is arbitrary, completely unmotivated, since it shares nothing with any other Bangla word. Notice that one is talking about items in lexical storage, not commenting on assembly.

The differently conceptualized pair "transparent/ opaque" does involve the assembly process. Transparency refers to the undistorted compositionality of an utterance. An utterance is compositional if no opaque barrier within it (such as a

world-creating predicate or modal operator) distorts or fragments the cumulation of part-interpretations assembling the interpretation of the whole. Such thinking views an utterance as consisting of constructions and ultimately as consisting of minimal utterances – independently usable words – and of those dependent words that have to work with them to build viable utterances.

We need some clarity about how the concept pair "arbitrary/ motivated" that helps us make sense of inter-sign relations in the lexicon is to be articulated vis-à-vis the differently conceptualized pair "transparent/ opaque" that has to do with sentence assembly. Merging the two into one concept pair is not a helpful move to make. But formalists experience a strong temptation – emanating from the residual structuralism in their thinking – to conflate the two. Why are they tempted to do this? And why should substantivists object?

The temptation goes back to structuralism, under whose assumptions relatively motivated signs like *kukurer* 'dog's' count as composite signs – without any interrogation of the nature of the process that composes these composite signs. A linguist who seeks maximal formal generality is thus liable to wish to fuse the two concept pairs along the following lines: – A language is anchored in a basic vocabulary consisting of simple (entirely arbitrary/ unmotivated) signs. Every relatively motivated sign is a composite sign, a construction composed of simple or composite signs. The patterns of the composition phenomena of a language are exhaustively describable in terms of rules. Rules specify opacity factor effects where necessary and implement transparent compositionality elsewhere.

We are focusing on the illegitimacy of such conflation of word complexity and sentence assembly. This leads us to portray the conflation as consistently appealing to the formal device "rule of grammar". Some readers are likely to conclude that we may thus be offering an obsolete picture of what the formalists are really doing. However, the distinction between descriptive rules and explanatory principles is hardly new. From its very inception, the formalist programme in generative grammar set itself the goal of examining the ways in which the abstract patterning that holds across rules lends itself to maximally general, "principled" characterization – termed "explanatory" rather than "descriptive" since the early sixties. Serious scrutiny of what the formalists have produced after the operational transition from rules to principles shows no discontinuity with the avowedly structuralist beginnings of transformational grammar. "Principle"-focused formalistic writing does indeed switch off such descriptive devices as construction-specific rules. But it retains the structuralist visualization of a sentence as an array of minimal bearers of meaning. That

picture postulates a hierarchy of arbitrary atomic units niched into substructures, these in turn niched into larger substructures, all the way up to the sentence itself.

Why should a substantivist have a problem with this? Does substantivism not recognize the fact that language involves wholes containing parts?

The substantivist's problem with this has to do with the role of novelty in sentence assembly. Sentences are freshly assembled. Words are the starting point from which assembly takes off. Motivation/ arbitrariness handles supplies. Transparency/ opacity is about assembly. A view that runs these together, in the name of a unified treatment of constituent-constitute relations, surely lets the generative revolution down – a revolution that was supposed to celebrate the constitutive novelty of sentence assembly.

Substantivism is about keeping faith with the core commitment of generative grammar, instead of celebrating the structuralist residue in our legacy. A unified treatment of constitute-constituent relations is an inappropriate generalization. A sentence is a fresh assembly, not a stale constitute.

In this section, we explore a few cases where the distinction between opacity/ transparency and arbitrariness/ motivation becomes an empirical matter. If the conflation of opacity with arbitrariness were valid, then there should be only one operative economy involved. But we find intersecting economies at work. To make sense of the data indicating this, we need to establish a division of labour between the two concept pairs.

Let us begin with irregular verbs, to take a cue from section 1. Consider the case of Bangla causatives, beginning with regular forms:

(13) (a)   rinaa duTo SaRi kacbe
              Rina two saris will.wash
              'Rina will wash two saris'
      (a')   jitu rinake diye duTo SaRi kacabe
              Jitu Rina by two saris will.make.wash
              'Jitu will make Rina wash two saris'
      (b)   korim tomake almari debe
              Korim you cupboard will.give
              'Karim will give you a cupboard'
      (b')   mOheS korimke diye tomake almari deWabe
              Mahesh Karim by you cupboard will.make.give
              'Mahesh will make Karim give you a cupboard'

We see in the examples given at (13) that the causative verbs of (13a', b') have a regular formal correspondence with the base verbs of (13a, b). The causative has an additional /(W)a/ within the verbal word. Some causatives are termed irregular because they do not match this template:

(14) (a) morle to Ek bari morbo
           if.die well one time.Emph will.die
           'If I die, well, I'll only die once'

      (a') prane marle (*mOrale) to Ek bari marbe (*mOrabe)
            life.Loc if.kill well one time.Emph will.kill
            'If they take my life, well, they'll kill me only once'

      (b) kaMcer baTi obhabe poRle to bhangbei
           glass bowl thus if.falls Prt will.break.Emph
           'A bowl made of glass, if it falls like that, will of course break'

      (b') kaMcer baTi tumi obhabe phelle (*pORale) to bhangbei
            glass bowl you thus if.drop Prt will.break.Emph
            'A bowl made of glass, if you drop it like that, will of course break'

      (c) eSOb rastaY gaRi Oto jore colbe na
           these roads.Loc car so fast will.go not
           'Cars won't go so fast on these roads'

      (c') eSOb rastaY tumi Oto jore gaRi calabe ki? (*cOlabe)
            these roads.Loc you so fast car will.drive Q
            'Will you drive so fast on these roads?'

      (d) tOrkariTa aSche
           the.vegetable is.coming
           'The vegetable is on its way'

      (d') ora tOrkariTa anche (*aSacche)
            they the.vegetable are.bringing
            'They are bringing the vegetables'

      (e) chatrira e ghOre thakuk
           the.girl.students this room.Loc let.stay
           'Let the girl students stay in this room'

      (e') ora chatrider e ghOre rakhuk (*thakak)
            they the.girl.students this room.Loc let.keep
            'Let them keep the girl students in this room'

      (f) oder meYe chOTaY ghum theke oThe
           their daughter at.six sleep from gets.up
           'Their daughter gets up at six'

(f')   ora meYeke chOTaY ghum theke tole (*oThaY)
        they daughter at.six sleep from rouse
        'They make their daughter get up at six'

As we see at (14a'-f'), the irregular causatives in the primed examples not only fail to match the base verbs (of the primeless examples (14a-f)) along the lines of (13), but actively prevent the regular causative counterparts from surfacing (we indicate this by presenting those forms and starring them).

However, there are special "sarcastic" contexts in which the regular causatives normally blocked by such irregular causatives make a cameo appearance – calling for theoretical commentary. Consider (14a"-f"), examples of the Sarcastic Causative:

(14") (a")   tumi bujhi morbe bhabcho? mOracchi!
                you Q will.die think? I'm.making.die!
                'You think you'll die, do you? I'll make you die!'
       (b")   Etogulo baTi poRe gElo, eTao poRbe bujhi? pORacchi!
                so.many bowls fell.& broke, this.too will.fall Q? I'm.making.fall!
                'So many bowls fell and broke, now it's this one's turn? I'll make it fall!'
       (c")   tomar gaRi eSOb rastaY cOle? cOlacchi!
                your car these roads.Loc go? I'm.making.go!
                'Your car goes on these roads, does it? I'll make it go!'
       (d")   OboSeSe Ekhón tOrkari aSche? aSacchi!
                at.last nów vegetable is.coming? I'm.making.come!
                'Now the vegetable arrives at last? I'll make it come!"
       (e")   chatrira e ghOre thakbe? thakacchi!
                girl.students this room.Loc will.stay? I'm.making.stay!
                'The girl students will stay in this room, will they? I'll make them stay!'
       (f")   meYeTa chOTaY ghum theke oThe? oThacchi!
                the.girl at.six sleep from gets.up? I'm.making.get.up!
                'The girl gets up at six, does she? I'll make her get up!'

In the pragmatically marked context exemplified at (14a"-f"), Bangla makes available Sarcastic Causative verbs, which carry a characteristic sarcasm-heavy intonation contour. These verbs flaunt precisely the regular causative templates

that Bangla takes plenty of trouble to avoid at (14a'-f'). What really is going on at (14a''-f'')?

A Sarcastic Causative form in Bangla is always a single, unbroken word. We would expect it – if arbitrariness and opacity were identical – to exhibit idiosyncrasy simply because it is a word rather than a phrase. But the sarcastic causative is manifestly as predictable as periphrastic causatives (as in *make him do it*), with which it turns out to share three properties that lexical causatives do not:

(15)   (a) phonological fidelity: the sarcastic causative mimics the base word
             closely;
        (b) semantic invariance: this causative ranges over all the uses of the
             base;
        (c) device independence: the mapping from the base onto this causative is
             consistent and does not diversify into different shapes for different
             base-causative dyads;

Irregular forms are said to block the regular forms one would have expected. But 'our' phenomenon, suspending the Blocking effect, exhibits what has been called Deblocking. En route to an explanatory analysis, it may help if we add another case of Deblocking – this time from English – to the basket of pertinent data:

(16)   (a)   life
        (a')   lives
        (b)   wife
        (b')   wives
        (c)   knife
        (c')   knives
        (aa)   Life
        (aa'') Lifes

The irregular plurals *lives, wives, knives* face no competition in English in default contexts, where they routinely block the regular plurals *lifes, wifes, knifes*. However, when we say *Life* – the name of a popular magazine from the fifties – some sort of anti-irregular context seems to get switched on. The regular *Lifes* surfaces here, unblocked by the irregular plural.

It is easy to satisfy ourselves that the association of anti-irregular plurals with the proper noun niche is a robust phenomenon. The French proper noun *Ciel* 'Sky' also exhibits the anti-irregular plural *Ciels*. In contrast, when the word is used as a common noun, the irregular plural *cieux* 'skies' always blocks *ciels*.

Tentatively, we propose the following descriptive summary of the anti-irregularity facts. In certain pragmatically marked contexts, a syntactic specification such as 'causative' is able to elicit forms exemplifying the most general plural/ causative templates that the morphology of the language concerned provides. Note that the phenomenon does not merely set aside normal irregular causatives in favour of otherwise unobserved regular causative schema instantiations. It even elicits causatives of verbs that otherwise resist the causativization process. For instance, the word for 'cause to sneeze' is marginal in ordinary Bangla. Sentence (17a) verges on ungrammaticality. The normal way to convey the sense of (17a) is (17b). But the sarcastic context brings even that verb into currency, as in (18):

(17)   (a) ???ei mOSlaTa SObaikei haMcaY
              this spice everybody makes.sneeze
           'This spice makes everybody sneeze'
         (b) ei mOSlaTate SObari haMci aSe
              this spice.Loc everybody.Gen sneezes come
           'Given this spice, everybody sneezes'

(18)   sophar nice lukiye tumi haMcbe? haMcacchi!
         sofa under hiding you will.sneeze? I'm.making.sneeze!
         'You hide under the sofa and sneeze, do you? I'll make you sneeze!'

Likewise, the Sarcastic Passive, as in (19a), also elicits impersonal passives of verbs normally resistant to passivization, as in (19b):

(19)   (a) abar chobi aMka hocche!
              Prt picture drawing Aux
           'And now (unspecified agent) is painting, is s/he!'
         (b) abar puliSer hate dhOra pORa hocche!
              Prt   police by caught getting Aux
           'And now (unspecified agent) is getting caught by the police, is s/he!'

What gives anti-irregular morphology such unimpeded access to forms otherwise unavailable in the lexicon?

What gives rise to these special effects, we suggest, is a bypassing of the lexicon. One way to implement this idea is as follows, in the case of the anti-irregular causative in Bangla:

(20) $[\text{Xe}]_{\text{T, PresSimp, 3p}} \longleftrightarrow [\text{Xacchi}]_{\text{T, PresProg, Caus, 1p}}$

Probal Dasgupta

This syntactically deployed Word Formation Strategy specifies the two sides in terms of the syntactic node T[ense] rather than the lexical category V[erb] and thus induces word formation on line, during syntactic tree assembly. For a similar analysis of the anti-irregular plural in English, see Dasgupta (2003: 69), where these forms are termed 'transparent'.

We now wish to add that these anti-irregular formations are indeed 'transparent' in the sense that they reflect UG processes capable of bypassing the 'arbitrary' workings of the lexicon even when they generate a word. The material for which section 1 proposed 'wild card' solutions was unusually irregular and pertained to arbitrary lexical storage. In contrast, the phenomenon we are now looking at is unusually regular and has to do with transparent (syntactic) assembly.

This contrast between the storage of arbitrary material and the fresh assembly of transparently compositional utterances indicates the presence of two formally distinct economies. What needs to be provided is an empirical demonstration that these economies not only coexist but intersect. For this purpose we turn to nominal classification in Bangla.

Rehearsing some facts about nominal classification phenomena may help set the stage for this part of our exposition. Bangla displays noun classification phenomena that invite description, as the following examples involving numerals indicate, in terms of classification formats, not distinct 'classifiers' (Dasgupta & Ghosh 2007, Dasgupta 2008):

(21) a. *EkTa   meye   jabe.*   b. *Ekjon   meye   jabe.*
         one.Gnl   girl   go.Fut   one.Hum girl   go.Fut
         'One girl will go.'   'One girl will go.'

(22) a. *duTo   meye   jabe.*   b. *dujon   meye   jabe.*
         two.Gnl   girl   go.Fut   two.Hum girl   go.Fut
         'Two girls will go.'   'Two girls will go.'

(23) a. *duTo ghOr khali ache.*   b. **dujon   ghOr*
         two.Gnl room vacant Cop   two.Hum room
         'Two rooms are vacant.'

(24) a. **duTo   bhOdromohila*
          two.Gnl   lady
      b. *dujon   bhOdromohila jaben.*
          two.Hum   lady   go.Fut.Hon
          'Two ladies will go.'

The noun *meye* 'girl' can occur in Bangla either with a general numeral such as *EkTa* 'one.Gnl', *duTo* 'two.Gnl', or with a human numeral such as *Ekjon*

'one.Hum', *dujon* 'two.Hum'. Some nouns are more selective. Thus, *ghOr* 'room' cannot take a human numeral, hence the starred status of (23b). In contrast, *bhOdromohila* 'lady' never occurs with a general numeral, which is why (24a) is ill-formed. A detail that will matter later in the discussion appears at the verb: the honorific future form *jaben* 'will go' at (24b) contrasts with the default or non-honorific *jabe*.

Classification formats are also available at Det, as in (25), or N, as in (26) (the glosses NuanIndiv and NuanColl, for nuanced individuation and nuanced collectivity, label poorly understood feature matrices):

(25)   a. *konTa*   b. *konjon*
        which.Gnl   which.Hum
        'which one?'   'which one?'
        c. *konkhana*   d. *kongulo*
        which.Inan   which.Coll
        'which one?' (inanimate)   'which ones?'
        e. *konTi*   f. *konguli*
        which.NuanIndiv   which.NuanColl
        'which one?'   'which ones?'

(26)   *ei*   *meyeTa*
       this girl.Gnl
       'this girl'

N and Det cannot, however, compete with Numerals and other Quantifiers as far as classificatory richness is concerned:

(27)   a. *EkTa   deyal*   b. *Ekjon   bhOdrolok*
        one.Gnl wall   one.Hum   gentleman
        'a wall'   'a gentleman'
        c. *Ekkhana   camoc*   d. *kOtokgulo   ciruni*
        one.Inan   spoon   couple.Coll comb
        'a spoon'   'a couple of combs'
        e. *kOyekTi   gan*   f. *Onekguli   gan*
        a.few.NuanIndiv song   many.NuanColl   song
        'a few songs'   'many songs'
        g. *EtoTuku   ca*   h. *Ekgacha   laThi*
        this.much.Dimin tea   one.Quirky stick
        'so little tea'   'a (walking) stick'

i. *EkpaTi   juto*
   one.Quirky shoe
   'a shoe'

j. *Onekkhani   SomOy*
   much.Expanse time
   'a lot of time'

k. *EtoTa   kOfi*
   this.much.Extent   coffee
   'so much coffee'

The Det site does not license such forms as **kongacha, *konpaTi, *konkhani* 'which.Quirky, which.Quirky, which.Expanse' (where 'Quirky' signals the extremely specialized lexical selection associating certain classification formats with certain nouns) and cannot switch on the 'Extent' meaning observed at (27k) to enable the *konTa* of (25a) to carry the additional sense 'what quantity'.

This exercise helps choose between a 'classifier morpheme' view of the matter and the WWM approach on empirical grounds. A morpheme-based analysis must assign clear and distinct feature compositions separately to a Det/ Num/ Q/ N base and to a Classifier affix morpheme. It has been shown (Dasgupta 2007) that even the simple grouping of the common 'Classifier morphs' into 'Classifier morphemes' is an unfeasible project.

Although Dasgupta (2007) provisionally accepted, for argument's sake, the working assumption that a 'Classifier morph' can be separated from a 'base' in featural terms, it has been shown (Dasgupta & Ghosh 2007) that even at the feature matrix level the separability assumption is unsustainable. What classification formats can a noun exhibit in order to mark definite/ specific readings (specific with a demonstrative and definite elsewhere)? No noun appears in a human /Xjon/ format, as shown in (28), but the /XTa/ format is widely used for singulars and /Xgulo/ for plurals, see (29), while /Xkhana/ marks inanimate singulars, as in (30):

(28)   a. **meyejon*
           girl.Hum
           'the girl'

        b. **bhOdromohilajon*
           lady.Hum
           'the lady'

        c. **upacarjojon*
           vice-chancellor.Hum
           'the vice-chancellor'

(29)   a. *meyeTa, *meyekhana*
           girl.Gnl, *girl.Inan
           'the girl'

        b. *meyegulo*
           girl.Coll
           'the girls'

        c. *ei   meyeTa*
           this girl.Gnl
           'this girl'

        d. *ei   meyegulo*
           this girl.Coll
           'these girls'

(30)   a. *camocTa,   camockhana*   b. *camocgulo*
         spoon.Gnl, spoon.Inan   spoon.Coll
        'the spoon'   'the spoons'
     c. *ei camocTa/ camockhana*   d. *ei   camocgulo*
         this spoon.Gnl/ spoon.Inan   this spoon.Coll
         'this spoon'   'these spoons'
     e. *DimTa,   Dimkhana*   f. *Dimgulo*
         egg.Gnl,   egg.Inan   egg.Coll
         'the egg'   'the eggs'
     g. *ei   DimTa/ Dimkhana*   h. *ei   Dimgulo*
         this   egg.Gnl/ egg.Inan   this egg.Coll
         'this egg'   'these eggs'

Some readers will need to unscramble this by examining the interaction between classification format exponence and the noun, the numeral/ quantifier and the determiner; one account of that traffic is provided in Dasgupta & Ghosh (2007). To summarize, a single Bangla nominal structure can carry classification features at no more than one of the sites Det, Num/Q, and N.

      We turn now to issues related to the way verbs agree with nominals for honorificity (recall the discussion of (24b)). The phrase 'these five students' translates two different Bangla phrases:

(31)   a. *ei   paMcjon chatro*   b. *ei   paMcTa chatro*
        this five.Hum student   this five.Gnl student

The numeral *paMcjon* in (31a), positively specified for a Human feature, contrasts with (31b)'s numeral *paMcTa*, whose feature composition is maximally general. Numerals appear either skeletally, when we count *Ek dui tin car paMc* 'one two three four five', or in this format that carries classification features. The present analysis describes (31a,b) in terms of the Word Formation Strategies (32) and (33):

(32) WFS for Human Numerals
     $[X]_{Num}$ ←→ $[Xjon]_{Num, Cla, Hum}$

(33) WFS for General Numerals
     $[X]_{Num}$ ←→ $[XTa]_{Num, Cla, Gnl}$

Bangla verbs agree with their subject for Person and Honorificity. While a pronoun, as in (34), must formally commit itself to an Honorificity value

Probal Dasgupta

(Intimate, nonHon, or Hon), a noun is, within limits, free to refer to individuals of varying degrees of honour, as shown in (35):

(34) 'You will go tomorrow', three variants:
    a. *tui   kal   jabi.*
        you.Intim tomorrow go.Fut.Intim
    b. *tumi   kal   jabe.*
        you.nHon tomorrow go.Fut.nHon
    c. *apni   kal   jaben.*
        you.Hon   tomorrow go.Fut.Hon

(35) 'My student will go tomorrow', two variants:
    a. *amar chatro kal   jabe.*
        my   student tomorrow go.Fut.3p.nHon
    b. *amar chatro kal   jaben.*
        my   student tomorrow go.Fut.Hon

Grammatically, any noun can take either Hon or nonHon agreement. When the noun means 'baby' or 'goat', Hon agreement signals irony. If the noun means 'president' or 'queen', nonHon agreement indicates a speaker's intention of expressing disrespect. Such deviations do not jeopardize grammaticality. Pronouns, however, trigger feature-driven agreement. If this requirement is not met, the results are neither ironic nor disrespectful, but sharply ungrammatical. Compare (34) with the following:

(36) **tumi   kal   jaben.*
      you.nHon tomorrow go.Fut.Hon
(37) **apni   kal   jabe.*
      you.Hon tomorrow go.Fut.nHon

Particular nouns have no lexically specified absolute Hon values. Formally the freely assigned Hon value a given nominal phrase carries triggers agreement. Does the noun control this Hon value?

(38) a. *ei   paMcjon   chatro kal   jabe.*
        this   five.Hum   student tomorrow   go.Fut.3p.nHon
    b. *ei   paMcjon   chatro kal   jaben.*
        this   five.Hum   student tomorrow   go.Fut.3p.Hon
        'These five students will go tomorrow.'
(39) a. *ei   paMcTa   chatro kal   jabe.*
        this   five.Gnl   student tomorrow   go.Fut.3p.nHon
        'These five students will go tomorrow.'

b.	*ei	paMcTa	chatro	kal	jaben.
	this	five.Gnl	student	tomorrow	go.Fut.3p.Hon

Human classification features carried by the numeral are compatible with both nonHon and Hon agreement, we find at (38a,b). But (39) shows that a General numeral triggers nonHon agreement, sharply excluding Hon.

What form should the proper description of this contrast take? We have seen at (35) that a noun can in principle accept both values of Hon. That (38b) contrasts with (39b) shows that the General feature matrix resists Hon agreement, whereas the Human feature composition of the nominal phrase in (38b) permits it. Just what needs to be said to describe both this fact and the rigidity of pronouns?

We refer the reader to a detailed study by Ghosh (2006) and, for our purposes, consider only the properties of one example, built around *upacarjo* 'vice-chancellor (university president, rector)':

(40)   a. ??*ei dujon   upacarjo   kal   jabe.*
          this two.Hum   VC   tomorrow go.Fut.3p.nHon
          'These two vice-chancellors will go tomorrow.'
       b. *ei   dujon   upacarjo   kal   jaben.*
          this two.Hum   VC   tomorrow go.Fut.3p.Hon
       'These two vice-chancellors will go tomorrow.'

(41)   a. *ei   duTo   upacarjo   kal   jabe.*
          this two.Gnl   VC   tomorrow go.Fut.3p.nHon
        'These two vice-chancellors will go tomorrow.'
       b. **ei   duTo   upacarjo   kal   jaben.*
          this two.Gnl   VC   tomorrow go.Fut.3p.Hon

While *chatro* 'student' is a neutral noun, *upacarjo* 'vice-chancellor' carries an Hon expectation, which (40b) meets. If a speaker intends disrespect, one expects the offensive use of language to go all the way, as in (41a), where the General (rather than Human) features of the numeral 'two' and the nonHon agreement on the verb 'will go' both express disrespect. (40a) is so puzzling as to sound like an error; the choice of the Human format for the numeral, together with the pragmatic default of honour for vice-chancellors, leads us to expect an Hon verb, but we get the puzzlingly nonHon verb *jabe* instead. To rescue (40a), we can imagine it to be uttered by a senior figure who is so far above all vice-chancellors in status that s/he can afford to use nonHon verbs for them casually, but who

Probal Dasgupta

wishes to avoid blatant disrespect and thus sticks to the Human format. The need to imagine such a special viewpoint for (40a) is what makes it nearly uninterpretable.

Now, suppose you are a disrespectful speaker and would use (41a). You would then say 'The vice-chancellor will go tomorrow' in the singular as (42). If you wish to show normal respect as in (40b), however, your choices are (43a,b). The Nuanced Individuation form, (43a), carries mild irony. The unformatted noun in (43b) can be diagnosed as a case of UG imposing a transparent stopgap in a niche left unoccupied by the arbitrary logic of classification formats in the particular grammar of Bangla. For some evidence for our diagnosis, note that even mild pejoration at *duTi* contradicts honour at the verb so severely as to nearly star (44a):

(42)   *upacarjoTa*   *kal*   *jabe.*
              VC.Gnl   tomorrow   go.Fut.3p.nHon

(43)   a.   *upacarjoTi*   *kal*   *jaben.*
              VC.NuanIndiv tomorrow   go.Fut.3p.Hon
        b.   *upacarjo*   *kal*   *jaben.*
              VC   tomorrow   go.Fut.3p.Hon

(44)   a.   ??*upacarjo-duTi*   *kal*   *jaben.*
              VC-two.NuanIndiv   tomorrow   go.Fut.3p.Hon
        b.   *upacarjo-dujon*   *kal*   *jaben.*
              VC-two.Hum   tomorrow   go.Fut.3p.Hon
        c.   *upacarjo-duTo*   *kal*   *jabe.*
              VC-two.Gnl   tomorrow   go.Fut.3p.nHon
              'The two vice-chancellors will go tomorrow.'

In other words, speakers choose between the respect-preserving default (44b) and the overtly disrespectful alternative (44c), and have no use for (44a), except perhaps to convey extreme irony. But (43b) sounds normal and (43a) comes out as an only slightly ironic variant. What does this indicate?

Our reading is that the UG default at (43b) and Bangla's Nuanced Individuation format /NTi/ at (43a) must be stepping in to fill a language-particular system gap. Revisiting (28-c) helps identify the gap in question (the format /Njon/ is starred); but the point of interest is the availability of two fillers, not the gap itself. The language-particular system offers a limited extension of /NTi/ at (43a) (limited in that (44) makes /NumTi/'s non-participation evident); UG offers the option that an unformatted Hon noun can take on the definiteness

features (see Dasgupta & Ghosh 2007 for details) normally associated with a classification formatted noun; neither of the fillers blocks the other.

We now need to work out just how the arbitrary, word-carried particular realities of a language interact with UG.

More concretely, the question is: how are we to make sense of the fact that, in a context where resources have to be stretched to meet unusual needs, what the arbitrary or particular face of Bangla has to offer by way of a minimal stretching of the logic of classification formats in order to fill the gap noted at (28) – namely, form (43a) – neither blocks nor gets blocked by (43b), the form that the transparent or UG-welcoming face of the language offers as a filler for the same gap?

To see that this fact contradicts the formalist doctrines, we may need to first imagine, along formalist lines, a single dimension ranging from "most arbitrary/ opaque" to "most motivated/ transparent", conflating the concept pair "arbitrary/ motivated" with the pair "opaque/ transparent". We then see that such a fusion of opacity with arbitrariness entails the prediction that, of the two choices (43a, b), one would block the other. A stronger, formalist-doctrine-focused variant of the prediction would say that only (43a) should count as well-formed, for (43b) flies in the face of the requirement in Bangla that a definite nominal should use a classification format in the nominal syntagm. Comparing this prediction, in either the weak or the strong variant, with the facts, we observe that (43) contradicts this prediction. The mutually unprevented availability of (43a, b) confirms the substantivist conjecture that provincial arbitrariness and transparency manage intersecting economies. Both of these economies, intersectingly, determine such matters as whether word A or word B is going to be available to a speaker of Bangla in context X or context Y.

The point that arbitrariness and transparency run distinct economies was made earlier (Dasgupta 2007: 170), in a context that focused on the consequences of the maximization of 'compact' arbitrariness in lexical storage and of the distinct maximization of transparency in the compositional assembly process of the syntax. In the present paper, we concentrate instead on the way these economies intersect, and thus on the logic of multiple validation. If language is driven by intersecting economies associated with lexical storage and with fresh assembly, it follows that parallel validation with respect to distinct sets of constraints is bound to raise formal issues. What do these issues look like on the ground? This is the question that comes to the fore in section 3.

Probal Dasgupta

3. Strategies and Shadows

Some advances in our understanding of validation in phonology and syntax have emerged from the study of ill-formed strings. What we have said about Word Formation Strategies so far concerns only the possible words that they serve to illuminate. But we need also to investigate what a strategy pushes into the shadow – its invalidations. To this end, consider WFS (45a), which interrelates the Bangla words shown at (45b-d):

(45) (a) $[X]_N \leftrightarrow [XWala]_N$, 'someone professionally concerned with X'
      (b) aiskrim 'icecream'   aiskrimWala   'icecream seller'
      (c) baRi 'house'   baRiWala   'landlord'
      (d) baMSi 'flute, pipe'   baMSiWala   'flautist, piper'
      (e) *baMSiWalaWala 'piper seller'

If we come across a vendor selling toy Pied Pipers of Hamelin, an unconstrained application of (45a) might make us call him a *baMSiWalaWala* 'a piper seller'; but this word, we note at (45e), is morphologically unavailable. Likewise, English does not enable the use of *flautistist* to describe a social scientist who studies the category of flautists.

    We propose to call this phenomenon the 'strategy shadow' cast by a WFS. Along the same lines, (46a)-(48a) in Bangla enable (b, c) in each set, but disable the strategy shadow forms at (d, e):

(46) (a) $[X]_N \leftrightarrow [Xoj]_{A/N}$ '(something) originating from X'
      (b) jOl   'water'   jOloj   'water-born (organism)'
      (c) bon   'forest'   bonoj   'forest-produced/ product'
      (d) jOloj (see (b))   *jOlojoj 'aquan-born'
      (e) bonoj (see (c))   *bonojoj 'sylvan-born'

(47) (a) $[X]_N \leftrightarrow [Xhin]_{A/N}$ '(someone) lacking X'
      (b) griho 'home'   grihohin '(someone) homeless'
      (c) bitto 'wealth'   bittohin '(someone) penniless'
      (d) grihohin (see (b))   *grihohinhin 'homelessless'
      (e) bittohin (see (c))   *bittohinhin 'pennilessless'

(48) (a) $[X]_N \leftrightarrow [Xbhoji]_{A/N}$ '(organism) feeding on X'
      (b) trino 'grass'   trinobhoji 'grass-eating/ eater'
      (c) pipilika 'ant'   pipilikabhoji 'ant-eating/ eater'
      (d) trinobhoji (see (b))   *trinobhojibhoji 'grass-eater-eater'
      (e) pipilikabhoji (see (c)) *pipilikabhojibhoji 'ant-eater-eater'

The fact that WFSs in general invalidate double application products needs to be contrasted with the different behaviour of freshly assembled phrases. A minimal pair useful for this purpose is available. At (50d, e) one finds the syntactic means to express the notions that the morphology is unable, at (49d, e), to format as single words – and we should add that by superimposing (50d, e)'s contrastive stress on (49d, e) does not remove their ill-formedness:

(49) (a) [X]$_N$ (([Xantor]N 'another X'
      (b) deS   'country'   deSantor   'another country'
      (c) gram 'village'   gramantor   'another village'
      (d) *deSantorantor 'another other country'
      (e) *gramantorantor 'another other village'

(50) (b) onno EkTa deS
           other one country
           'another country'
      (c) onno EkTa gram
           other one village
           'another village'
      (d) ónno EkTa onno EkTa deS
           other one other one country
           'anóther other country'
      (e) ónno EkTa onno EkTa gram
           other one other one village
           'anóther other village'

The reason that (50) goes through is that the syntax performs fresh and free assembly. (49d, e) fail because a strategy cannot operate in its own shadow – a phenomenon that calls for a formal account.

      But is it indeed the case that every WFS casts a shadow systematically invalidating what the double application of the strategy would have produced? Do (51) and (52) from Bangla and (53) and (54) from English not counterexemplify the claim that this phenomenon is perfectly general?

(51) (a) pitamOho 'grandfather'
      (b) propitamOho 'great-grandfather'
      (c) propropitamOho 'great-great-grandfather'

(52) (a)   SOmaj 'society'
      (b)   SOmajbirodhi '(an) anti-social (element)'
      (c)   SOmajbirodhibirodhi '(an) anti-anti-social (element)'

(53) (a)   communist
      (b)   anti-communist
      (c)   anti-anti-communist

(54) (a)   language
      (b)   meta-language
      (c)   meta-meta-language

We comment on these cases in Appendix 1, in order to preserve the flow of this discussion and to make it easy for readers to develop their own solutions. Setting these aside, we now propose a general analysis of (45)-(50). Strategy (45a) as stated –

(45)   (a) $[X]_N \longleftrightarrow [XWala]_N$, 'someone professionally concerned with X'

– does not say only that one can move from *X* to the schema specified phonically as *XWala* and semantically as 'an X-concerned person'. Its bidirectional arrow also lets a user move, phono-semantically, from *XWala* to *X*. A foreign learner of Bangla who has not heard *kulpi* 'coolfi, a cold sweet', on hearing *kulpiWala* 'coolfi seller', will infer that *kulpi* is the word for whatever such a person is professionally concerned with.

Notice, then, that a WFS works, with reference to some set of paired examples underwriting the strategy – for instance, (45b-d) – by applying the bischematic template of (45a) to a word that fits either the *X* schema or the *XWala* schema. On the basis of this template matching, the strategy's action proceeds either right to left, adding *Wala* plus its semantics to an *X* that lacks it, or left to right, subtracting *Wala*-plus-semantics from an *XWala* that has it. A strategy is a toggle switch. In its bischematic design, the specification of *XWala* on the right-hand side implies that the *X* on the left does **not** have a *Wala*-plus-semantics in it. Applying this strategy to an *XWala* form yields an *X* form that involves reversing the 'add *Wala*' instruction, i.e. subtracting *Wala*-plus-semantics. There is no way to obtain an *XWalaWala* form by applying (45a).

In other words, the formal operation of word formation strategies itself entails the strategy shadow effect as a theorem.

Does this account invite comparison with Aronoff's (1976: 95-97) discussion based on Isačenko (1972) on what they both took to be "truncation rules which prevent surface suffix doubling"? Or perhaps with Aronoff's (1976:

37n4) observation that "Systematically, *-ly* does not attach to adjectives which themselves end in *-ly* (*silly/*sillily*)"? The status of Aronoff's truncation rules or of his constraint on *-ly* seems not to have been elaborated into a full-blown formalist analysis of a "suffix doubling" filter.

In order to take the debate further – regardless of particular approaches – morphologists will need to consider the possibility of paradigmatically relating the strategy shadow phenomenon to reduplication. To see the point, imagine that language design were to let word formation strategies apply iteratively and to produce 'accidental reduplication'. Such traffic would get in the way of reduplication existing as a distinct phenomenon. But natural language seems to have some use for reduplication (for some recent serious work on reduplication, see Singh 2005, Montaut 2008). It follows that reduplication needs space, and therefore must be visible. Thus language design must have features guaranteeing the non-generation of accidental reduplicants – features such as strategy shadow, if our account of the phenomenon is on the right track.

Whether we are on the right track is something we can check by triangulating – by asking whether strategies other than WFSs cast a shadow. At this juncture readers need to acquaint themselves with substantivist proposals for the proper treatment of arbitrariness in lexico-phrasal storage.

The most arbitrary material takes the form of words. Substantivist lexical entries are connected by Word Formation Strategies, already exhibited in this paper. Within this realm, other schools of thought formally demarcate degrees of arbitrariness by using either lexical strata or the word boundary/ morpheme boundary distinction. But substantivism adopts the WWM working hypothesis that there are no morpheme boundaries, and that postulating a word-internal word boundary, as in Aronoff (1976: 121-9) on the 'productive affix *#able*', is not a descriptively adequate solution. It is perhaps only fair to specify what our take is on these matters.

At the formal level, analyzing the material within WWM is a straightforward exercise. Those 'unproductive *+able/ +ible*' words that appear in pairs like *perceptive, perceptible, suggestive, suggestible, division, divisible, derision, derisible* are amenable to highly specified WFSs like *Xtiv~Xtibl, Xžən~Xzibl*, etc. But Aronoff's 'productive affix *#able*' corresponds to a simple, general WFS *X~Xəbl*.

Where does WWM or substantivism say this distinction should predict that the intricately arbitrary 'unproductive' cases shall contrast with the more iconic phonology and semantics of the 'productive' ones? WWM does not say; it

excludes the matter from the morphology. Substantivism comments that (non-) iconicity properties of words – and the contrast between general WFSs, which by preserving shape maximize iconicity, and highly specified WFSs, which do not – are handled by semiotics. Natural language words, as well as images and other objects of semiotic inquiry, invite semiotic description and explanation, in addition to formal linguistic analysis. This is what multiple validation is all about. Inter-word paradigmatic relations hold in the space of discourse, a domain that semiotic analysis has long claimed for its own. (Readers unfamiliar with the place of semiotics in substantivist theory may wish to revisit Dasgupta, Ford & Singh (2000: 177).)

What substantivists object to is not the decision by other linguists to present some semiotic results – in terms of strata, distinct boundaries, or other devices – within what they package as morphology, for this is a question of nomenclature. We object to the claim that these devices are continuous with the formal mechanisms of syntactic assembly. And it is that claim that lies at the heart of the formalist programme.

It is important to demarcate the lexical domain of arbitrariness from the process of fresh assembly in the sentence. If the relations between a word and its neighbours were subject only to the laws of the syntax, the task would be simple, and the substantivist account of lexico-phrasal storage would refrain from comment. But linguists recognize the special relations of a 'clitic' with its 'host' word, and constructions that in various ways elude syntactic generalization. Substantivist work has accordingly made formal proposals for the adequate treatment of the lexico-phrasal specification of these properties.

To handle clitics, we have proposed Word Extension Strategies, WESs (Dasgupta 2005: 61). Although WESs do not formally define the notion 'clitic' – this reticence is akin to what prevents WWM from formally registering the semiotically distinguishable degrees of arbitrariness – they in effect postulate one WES per clitic. In this paper, we have nothing to say about Word Extension Strategies.

In order to handle constructions that either resist general treatment in the syntax or correspond so directly to morphological devices that the continuity with them needs to be formalized, we have proposed Phrase Formation Strategies (Dasgupta, Ford & Singh 2000: 171).[1] To return to the main thread, in our bid to

---

[1] The formulation of PFSs by Dasgupta, Ford & Singh (2000: 171) stands, but the "WFS" (91b) at 2000: 172 would be formalized today as a Word Extension Strategy. Our "WFS" there gets the phonology wrong – it wrongly predicts /bhulbabe/ with a short [u], whereas the correct output is /bhul#bhabe/ with a phonetically long [u:].

Discourse and Generative Grammar: a substantivist approach

check whether our strategy shadow proposals are on the right track, it is to Phrase Formation Strategies (PFSs) that we now turn.

In addition to adjectival comparison WFSs, English also has PFSs introducing the comparative functor *more* and the superlative functor *most*. The formulations provided below abstract away from syntactic framework-laden details that will need to appear in any fleshed-out version of strategies (55a, b):

(55) (a) $[X]_{Adj} \rightarrow [more \ [X]_{Adj}]_{Compv}$
(b) $[X]_{Adj} \rightarrow [most \ [X]_{Adj}]_{Superlv}$

The comparative WFS in English associates light adjectives with comparative adjective words like *higher, lower, brighter*. In contrast, PFS (55a) associates heavier adjectives like *strenuous, intelligent, effective* with phrasal comparatives like *more strenuous*. The question now is whether (55a) casts a strategy shadow.

The logic of comparison – given the grammaticality of (56c) – would lead us to expect (56d) to be fine, but it is in fact ungrammatical, and this looks like a strategy shadow fact:

(56) (a) A is more effective than B (as a manager)
(b) C is more effective than D
(c) A is as much more effective than B as C is more effective than D
(d) *A is more more effective than B than C is more effective than D

Unless some other account of the ungrammaticality of (56d) is shown to be more persuasive, data set (56) stands as evidence for the claim that PFS (55a) is the right analysis of phrasal comparison of English, that it casts a shadow, and that (56d) falls within this shadow.

At this stage, issues of intermodular traffic arise. Compare (56) with (57) – we omit the (a)-(c) examples here to save space:

(57) (d) *A is more taller than B than C is taller than D

Under substantivist assumptions, a word such as *taller* does not in any sense arise from a syntactic structure of the *er MUCH tall* type. Thus, the syntax per se cannot monitor interactions between the shadow of PFS (55a) and the shadow of the WFS responsible for *taller*. What, then, prevents (57d)?

The strategy shadow phenomenon has to do with what the strategy is adding/ subtracting – here the Comparative specification, which actual instantiations must systematically lack in order to match the left-hand side schema in (55a) and thus to qualify as acceptable left-hand side input for that

Probal Dasgupta

strategy. (57d) then fails because *taller*, though an Adj, fails to match this feature of the schema, since it bears Comparative feature/s. The syntactic composition of the Comparative feature submatrix is where the action is. This submatrix, shared by PFS (55a) with the WFS that handles light adjective comparatives like *taller*, ensures that (57d) will not be generated through the action of (55a).

For this account to work, it is essential to distinguish syntax from semantics. Consider (58) –

(58)   (d)   ?John is more senior to Bill than Susie is to Mary

– which is, if not perfect for all speakers, clearly far more acceptable than (57d). If semantics and syntax were closely matched, we would expect the adjective *senior* to be a lexically irregular comparative (given secondary licensing by the morphology as a comparative), and we would expect (58d) to be just as bad as (57d). This expectation is not met. We conclude that the lexical item *senior* does not carry the syntactic feature submatrix that specifies true comparatives, even though its semantics must be extremely close to that of a true syntactic comparative such as *older*.

We cannot conclude, however, that the semantics per se has no direct bearing on strategy shadow phenomena. Consider (59a-h) from Bangla and (60) from English:

(59)   (a)   bhu   'earth'
        (a')   bhutOtto   'geology'
        (b)   SOmaj   'society'
        (b')   SOmajtOtto   'sociology'
        (c)   cikitSa   'treatment'
        (c')   cikitSaSastro   'the science of medicine'
        (d)   gonit   'mathematics'
        (d')   gonitSastro   'the study of mathematics'
        (e)   *bhutOttoSastro   'geology-science'
        (f)   *SOmajtOttoSastro   'sociology-science'
        (g)   *cikitSaSastrotOtto   'medicine-ology'
        (h)   *gonitSastrotOtto   'mathematics-ology'

(60)   (a)   *geographology, *geologography, *economicology, etc.
        (b)   *driverist, *chauffeurist, *conductorist, etc.

The unavailability of double marking of discipline status observed at (59e-h) in Bangla and (60a) in English, and of double marking of professional status observed at (60b) in English, suggests that there is a specific semantic factor to

reckon with. Next to cases of the strategy shadow phenomenon itself, such as *Xerer, *Xistist, *Xographography,* the semantic clustering of certain strategies gives rise to a "strategy cluster shadow" phenomenon. The way it works is that iterative applications of any two strategies that are clustermates sound just as bad as iterative application of one single strategy.

What needs to be mapped with care is the differentiation of effects. Do these specifically semantic effects seem to parallel certain doubling avoidance effects that appear at the word formation level but arise from phonological factors, as in Aronoff's (1976:37n4) footnote cited above? Or does the semantics work in a more intimate association with the syntax? At the present moment in such inquiry, the best we can do is sharpen these questions and cast our empirical net wider.

When parallel validation – with respect to different components of the language faculty – works smoothly, we have no way of detecting the separate contribution that the various components make to the process of validating a legitimate linguistic structure. The present section, focusing as it does on ill-formed cases, gives us some access to points at which the multiplicity of validation becomes visible, and the contributions of the various components fan out.

4. Questions about Constructions

When we approach the study of constructions that resist general treatment in the syntax, we face a frustratingly heterogeneous list of quirky phenomena. The Phrase Formation Strategy device – or devices from other frameworks that specify the properties of a syntactic construction – can provide initial coverage of some facts. But it is unclear how to advance our understanding of the niche that these distinctive phenomena occupy either in the process of sentence assembly, or in the realm of lexico-phrasal storage, or even in some intermediate zone. We thus need to focus on the problems that all inquiry in this domain must address. Recent years have seen the development of formalisms that manage the interface with the formal semantics, or that make it possible to harvest corpus data for phraseological listing. While these developments are welcome, the fundamental issues of the traffic between lexico-phrasal arbitrariness and syntactic transparency in these semi-transparent constructions do not come to the fore. It is to be hoped that the relatively unadventurous formal properties of PFSs will

make it possible for the substantivist framework to highlight questions that others are also going to have to face.

Consider some concrete cases from Bangla. The regular syntax of embedded non-finite constructions is exemplified at (61); the embedded verb appears in the infinitive form. At (62) we observe the conditional adverbial participle. The sentences in (63) show that these sequences are reversible in certain contexts.

(61)   (a)   tOndra gan gayte pare
               Tandra song sing.Inf can
               'Tandra can sing (songs)'
      (b)   bijon SaMtar kaTte ceYechilo
               Bijan swim Aux.Inf wanted
               'Bijan wanted to swim'
      (c)   robin projitke baRi jete debe na
               Rabin Prajit home go.Inf will.let Neg
               'Rabin won't let Prajit go home'

(62)   (a)   tOndra gan gayle projit cole jabe
               Tandra song sing.Cnd Prajit away will.go
               'If Tandra sings, Prajit will leave'
      (b)   projit baRi gele robin chOTphOT korbe
               Prajit home go.Cnd Rabin fidget will.do
               'If Prajit goes home, Rabin will fidget'

(63)   (a)   tOndra páre gan gayte
               Tandra cán song sing.Inf
               'Tandra cán sing (songs)'
      (b)   bijon céYechilo SaMtar kaTte
               Bijan wánted swim Aux.Inf
               'Bijan did want to swim'
      (c)   prójitke baRi jete robin debe na
               Prájit home go.Inf Rabin will.let Neg
               'Prájit going home is something Rabin won't allow'
      (d)   projit cóle jabe tOndra gan gayle
               Prajit awáy will.go Tandra song sing.Cnd
               'Prajit will léave if Tandra sings'
      (e)   robin chÓTphOT korbe projit baRi gele
               Rabin fídget will.do Prajit home go.Cnd
               'Rabin will fídget if Prajit goes home'

Discourse and Generative Grammar: a substantivist approach

We now consider some embedded conditional participles that call for PFS treatment:

(64)   (a)   projit baRi gele pare
               Prajit home go.Cnd can
               'It's best if Prajit goes home'
       (b)   tumi gan Sikhle paro
               you music learn.Cnd can
               'It's best if you learn music'

(65)   (a)   projit baRi gelé i pare
               Prajit home gó.Cnd Emph can
               'Prajit may as well go home, why doesn't he'
       (b)   tumi gan Sikhlé i paro
               you music léarn.Cnd Emph can
               'You may as well learn music, why don't you'

We would provide PFS formulations if their statements were of interest. But the point to look at – a point that PFS formulations cannot address – is that permutations are systematically excluded. Special intonation add-ons have no impact on (66a, b), and the presence or absence of the Emphatic particle makes no difference either:

(66)   (a)   *projit pare baRi gele (i)
               Prajit can home go.Cnd (Emph)
       (b)   *tumi paro gan Sikhle (i)
               you can music learn.Cnd (Emph)

Why should (66) be excluded? We take it that the conditional verbs in the special constructions of (64) and (65) are doing work normally reserved for infinitives. However, regular conditional adjuncts, as in (63d-e), and regular infinitival complements, as in (63a-c), are not averse to permutation. What then prevents permutation in (66a-b)? In what way does the construction, however specified, affect the fundamental properties of these verb forms in the syntax? The future can be fortified by sandwiching an Emphatic /i/ between two copies:

(67)   (a)   amra rajar SOngge dEkha korbo i korbo
               we king with meet will.do Emph will.do
               'We absolutely sháll meet the king'

Probal Dasgupta

(b)   o tomake churiTa pherot debe i debe
        s/he you the.knife back will.give Emph will.give
        'S/he definitely will give you the knife back'

This special construction allows several forms of permutation, but it excludes negation and interrogation. In contrast, the embedded conditional construction of (64)-(65) allows interrogation and negation. Suppose hit and miss methods enable us to work out just which regular features of the syntax are sidelined or cancelled in this or that special construction. Even if all those answers become available, how can we so dovetail the construction's special demands with the syntactic system's general traffic that our description will replicate just the right degree of traffic dislocation?

Our response is to appeal to the secondary licensing device proposed in section 1 and to look at the specific zone of clausal architecture that the construction targets. Secondary licensing locates the drama in the root sentence. The fortified future is a positive polarity construction and targets the zone where options of negation or interrogation would have surfaced; hence the unavailability of negation and interrogation in the case of (67). The embedded conditional participle targets a zone lower in the clausal architecture and thus leaves negation/ interrogation options unaffected, but freezes the non-finite plus finite sequence, whose constituents therefore cannot be separated.

Is this a principled response, though? Have we been able to effect a neat separation of lexical idiosyncrasy from constructional irregularity? Can we claim to have put all truly arbitrary material in a lexical box whose lid we know how to shut?

In order to conclude this study, we would like to point out that the lexical box does not have a lid we can shut. One type of lexical idiosyncrasy is described in terms of 'bound' words. Such a word is bound to a particular neighbourhood. For instance, in contemporary English, the bound words *betwixt* 'between' and *let* 'hindrance' occur only in the fixed locutions *betwixt and between* and *without let or hindrance*. One would imagine that bound words represent the peak of arbitrariness in natural language.

It is noted in Dasgupta (2006: 155-57) that a systematic class of bound words in Bangla are associated with a 'productive' WFS:

(68)   ki choTánTa i na chuTechi!
        what run.x Emph Particle I've.run
        'What a running I've run!'

(69)   ami ja ThOkán Thokechi tar juRi nei!
        I what cheat.x have.been.cheated its match isn't
        'Nobody can match the deception I've been through!'

(70)   lokTa amader ki bhogán bhugiyeche!
        the.man us what harassment has.harassed
        'What harassment that man has put us through!'

These words *choTan, ThOkan, bhogan* are bound – they always serve as cognate objects that must co-occur with a verb with which the WFS associates them. They carry a characteristic intonation contour expressive of frustration and must appear in a clause with an exclamatory wh-phrase. If we are able to refine sufficiently the formalism of PFSs, perhaps such data will lend itself to statement in terms of a WFS embedded in a PFS, though this lies beyond our current means.

To the extent that words bound hand and foot to a particular context count as especially arbitrary, there is something paradoxical about the fact that a languagewide pattern should be able to sponsor a systematic class of bound words – something oxymoronic about the expression "systematic class of bound words". B.N. Patnaik (personal communication) informs us that (68)-(70) are not unique to Bangla – that similar facts obtain in Oriya, a sister language.

We can get around the paradox, technically, by so defining the notion of 'bound word' that only a word obliged to co-occur with a specified neighbour shall count as bound. Exclamatory cognate objects, in that case, would then stop being bound – they are obliged only to co-occur with a specified **type** of neighbourhood, not with any specified neighbouring word or words. Nevertheless, it is surely odd that a language should produce an entire class of words only for use in such restricted contexts; surely no known theory of arbitrariness in natural language predicts such a phenomenon. To this extent, we should conclude that we do not yet know how to put a lid on the lexicon as a repository of arbitrariness.

One striking feature of the exclamatory cognate object phenomenon in Bangla pulls together some of the earlier strands in our discussion and helps bring our deliberations to a close. Not only does the phenomenon target the word level and thus trigger an unusual lexical mechanism – specifically, if we are on the right track, a WFS embedded in a PFS. The phenomenon also contributes to the formation of a syntactic exception of the type that attracts what we have called secondary licensing in the syntax, and therefore directly hits the syntactic

Probal Dasgupta

roof – it helps compose a root sentence. Unusual word type meets unusual clause type at a picnic of exceptions. For us, it is time to celebrate linguistic theory's ability to find space for a very wide range of facts within a moderately restrictive and well-understood theory of how the modules cooperate. We can look forward to more illumination – especially in the study of strategy shadow.

Appendix 1
We now revisit those examples from section 3 that appeared to be a problem for the strategy shadow hypothesis:

(51)   (a)   pitamOho 'grandfather'
       (b)   propitamOho 'great-grandfather'
       (c)   propropitamOho 'great-great-grandfather'

(52)   (a)   SOmaj 'society'
       (b)   SOmajbirodhi '(an) anti-social (element)'
       (c)   SOmajbirodhibirodhi '(an) anti-anti-social (element)'

(53)   (a)   communist
       (b)   anti-communist
       (c)   anti-anti-communist

(54)   (a)   language
       (b)   meta-language
       (c)   meta-meta-language

Readers will have come up with their own conjectures. We would hazard the guess that these examples show formal/ mathematical game-playing at work in language. It has long been known that academic users of a language deploying its verbal resources for formal/ mathematical purposes – such as mediaeval logicians reshaping Sanskrit to make the work of navya nyaaya ('the new logic') possible – routinely stretch these resources beyond what natural language use would permit. Their utterances, usually in the written mode, violate constraints that ordinary language use, outside the context of mathematical game-playing, consistently adheres to.

We realize of course that *anti-anti-communist* and *SOmajbirodhibirodhi* are not themselves words invented by mathematicians. We would nonetheless like to suggest that they are playful, constraint-violating imports into natural language from the formal/ mathematical realm. Specifically, we are claiming that the (c)-forms in (51)-(54) are not words obtainable by normal morphological means, but

loans from a special mathematical register of human activity that lies at the edge of language per se.

Some readers will jump to the conclusion that by making this move we are introducing an escape hatch that amounts to the destruction of falsifiability for our account. To preempt that jump, let us briefly point out that even in a language where much of the morphology is an explicitly mathematics-type exercise, the artificial language Esperanto, certain expected outcomes do not occur.

Esperanto allows the formation of words like *ŝafido* 'lamb' from *ŝafo* 'sheep', *kaprido* 'kid' from *kapro* 'goat'. Users of the language are creative and playful. Thus, one would have expected *filo* 'son' and *nepo* 'grandson' to have given rise to *nepido* 'great-grandson', *nepidido* 'great-great-grandson' and so on. But what one says in fact is *pranepo* for 'great-grandson' with the same *praX* device that appears in *praavo* for 'great-grandfather' based on *avo* 'grandfather'. Esperanto iterates *pra* the way English iterates *great*, but that is about all; it does not permit *socialismismo* for 'doctrinal attachment to socialism or *tajpististo* for 'someone who professionally deals with typists' or *lernigigi* for 'to cause someone to cause someone to learn' – although *instruigi* 'to cause someone to teach' is fine, indicating that the problem is formal rather than purely semantic. It would be a big mistake to imagine that the formal imagination, when left unfettered, does in fact run wild. It does not, and inquiry is needed to find out exactly what constraints it spontaneously observes.

One problem for the account has to do with double causativization, which several languages permit. That only causatives pose a problem indicates that the account as a whole is on the right track; but the problem of causatives remains unsolved.

Acknowledgements

Probal Dasgupta

References

Aronoff, Mark. 1976. Word Formation in Generative Grammar. Cambridge: MIT Press.

Aronoff, Mark. 1993. Morphology by itself: Stems and Inflectional Classes. Cambridge: MIT Press.

Chomsky, Noam; Halle, Morris. 1968. The Sound Pattern of English. New York: Harper and Row.

Dasgupta, Probal. 1977. The internal grammar of Bangla compound verbs. Indian Linguistics 38:2.68-85.

Dasgupta, Probal. 1989. Outgrowing Quine: towards substantivism in the theory of translation. International Journal of Translation 1:2.13-41.

Dasgupta, Probal. 1993. The Otherness of English: India's Auntie Tongue Syndrome. New Delhi/ London/ Thousand Oaks: Sage.

Dasgupta, Probal. 2001. On a vowel template asymmetry in Bangla verbs. Abbi, Anvita; Gupta, R.S.; Kidwai, Ayesha (eds) Linguistic Structure and Language Dynamics in South Asia: Papers from the Proceedings of the SALA XVIII Roundtable. Delhi: Motilal Banarsidass. 164-181.

Dasgupta, Probal. 2003. Antiopacity and Bangla causatives: a substantivist approach. Singh, Rajendra (ed.) Yearbook of South Asian Languages and Linguistics 2003. Berlin/ New York: Mouton de Gruyter. 47-70.

Dasgupta, Probal. 2005. Q-baa and Bangla clause structure. Singh, Rajendra, et al. (eds.) Yearbook of South Asian Languages and Linguistics 2005. Berlin/ New York: Mouton de Gruyter. 45-81.

Dasgupta, Probal. 2006. Bikiron, akoronpokkho ar kayapokkho ['productivity, formalism and substantivism']. Bangiya Sahitya Parishat Patrika 113:1-2.147-157.

Dasgupta, Probal. 2007. Advances in substantivist grammatical research. Nagaraja, K.S.; Bhatta, V.P.; Kulkarni-Joshi, Sonal; Pingle, P.M. (eds) Research Trends in Lexicography, Sanskrit and Linguistics: Proceedings of the Professor S.M. Katre Birth Centenary Seminar. Pune: Deccan College Research Institute.151-181.

Dasgupta, Probal. 2008. Transparency and arbitrariness in natural language. Singh, Rajendra (ed.) Annual Review of South Asian Languages and Linguistics 2008. Berlin/ New York: Mouton de Gruyter.3-19.

Dasgupta, Probal; Ford, Alan; Singh, Rajendra. 2000. After Etymology: towards a Substantivist Linguistics. Muenchen: Lincom Europa.

Discourse and Generative Grammar: a substantivist approach

Dasgupta, Probal; Ghosh, Rajat. 2007. The nominal left periphery in Bangla and Asamiya. Singh, Rajendra (ed.). Annual Review of South Asian Languages and Linguistics 2007. Berlin/ New York: Mouton de Gruyter. 3-29.

Ford, Alan; Singh, Rajendra; Martohardjono, Gita. 1997. Pace Panini. New York: Peter Lang.

Ghosh, Sanjukta. 2006. Honorificity-marking words of Bangla and Hindi: Classifiers or not? Bhashacintana 1.21-27.

Isačenko, A.V. 1972. Rolj usechenija b russkom slovoobrazovanii ['the role of truncation in Russian word formation']. International Journal of Slavic Linguistics and Poetics 15.95-125.

Montaut, Annie. 2008. Reduplication and 'echo words' in Hindi/ Urdu. Annual Review of South Asian Languages and Linguistics 2008. Berlin/ New York: Mouton de Gruyter. 21-61.

Singh, Rajendra. 2005. Reduplication in Modern Hindi and the theory of reduplication. Hurch, Bernhard (ed.) Studies in Reduplication. Berlin/ New York: Mouton de Gruyter. 263-81.

Singh, Rajendra, and Stanley Starosta (eds.) 2003. Explorations in Seamless Morphology. New Delhi/ London/ Thousand Oaks: Sage.

Two Types of Intervention Effects

Barry C.-Y. Yang
National United University, Taiwan

Abstract
The exploration on wh-intervention effects generally suffers from distributional variations both across languages and within a language. In this study I show that with evidence from Mandarin Chinese the puzzling variations of intervention effects can be easily sorted out. I propose that the variations are best captured if we categorize them into two types of syntactic effects, i.e., Minimality Effect and Competition Effect, each of which is triggered by independent factors.

1. Introduction

The exploration on *wh*-intervention effects generally suffers from distributional variations across languages. For instance, while *wh*-arguments in Korean and Japanese (1) are sensitive to intervention effects (Beck 1996, Beck and Kim 1997), they are not in Chinese (2) (Cheng & Rooryck 2002, Tsai 2004, Soh 2005).

(1) a. ??<u>Daremo-ga</u>   **nani-o**   yon-da-no   [Japanese]
         everyone-nom   what-acc read-past-Q
         'What did everyone read?'
     b. *<u>Daremo</u> **nani-o**   yom-ana-katta-no
         anyone   what-acc read-neg-past-Q
         'What did no one read?'

(2) a. {<u>Suoyoude/meige ren</u>}   dou mai **shenme**? [Chinese]
          all/every   person all   buy   what
          'What did all people/everyone buy ___?'
     b. {<u>Meiyouren/henshaoren/zuiduo liang-ge ren</u>} gan gen **shei** dajia?
          nobody/few.person/at.most two-Cl person   dare with who   fight
          'Who is the person x such that nobody/few people/at most two people
           dare(s) to fight with x?'   (from Soh 2005)

Barry C.-Y. Yang

On the other hand, even within a language, *e.g.*, English, an in-situ *wh*-argument is subject to positional variations with respect to intervention effects (Pesetsky 2000):

(3) a. *Which book <u>didn't</u> **which person** read ___?   (wh_1i-in-situ)
     b.   Which person <u>didn't</u> read **which book**?   (non-wh_1i-in-situ)

Meanwhile, the speaker judgment is also reported to be inconsistent (Lee 2001, Kuno & S.-Y. Kim 2004, Miyagawa & Endo 2004, among others).

Though recent studies tend to limit their discussions to the focus intervention effect (Kim 2002, 2005, Beck 2006) since it enjoys a stable distribution across languages, still language data show that it can be weakened in embedded contexts as in (4) (Tomioka 2007) or in D-linked ones (Miyagawa & Endo 2004).

(4) a. *<u>Taro-sika</u>   **nani-o**   yom-ana-katta-no   [Japanese, Tomioka 2007]
            -except what-acc read-neg-past-Q
         'What did no one but Taro read?'
     b. ?Kimi-wa [CP <u>Taro-sika</u>   **nani-o**   yom-ana-katta-to]   omotteiru-no?
         you-top   -except what-acc read-neg-past-comp think-Q
         'What do you think that no one read?'

This study shows that evidence from Mandarin Chinese provides a clear picture for us to sort out the puzzling variations of intervention effects. I propose that the variations can be best captured if we categorize them into two types of syntactic effects, i.e., Minimality Effect and Competition Effect, each of which is triggered by independent factors.

2. The Weak-Strong Divide

To begin with, Chinese exhibits a clear weak-strong divide with respect to intervention effects. That is, in non-focus context the intervention effect (Type-I) is weak so that only the *wh*-adverb is ruled out (5) whereas in focus context it (Type II) is so strong that even the *wh*-argument is ruled out (6).

(5) a. *Ta <u>bu/zongshi</u> renwei Lisi **weishenme/zenme** cizhi? (**wh*-adverb)
         he not/always think   Lisi why/how   resign
         'Why$_i$/How$_i$ does he not/always think Lisi resigned t_i?'
     b. Ta <u>bu/zongshi</u> renwei Lisi xihuan **shei**?   (^{ok}*wh*-argument)
         he not/always think   Lisi like   whom
         'Whom does he not/always think Lisi likes?'

(6)   *<u>Shi/Zhiyou</u> Lisi   chi-le   **shenme**?   (*_wh_-argument)
        be/only   Lisi   eat-Asp what
        Lit. 'What was x such that {it was Lisi who/only Lisi} ate x?'

In the spirit of Tsai (1994), the Type-I intervention in (5) is expected because the _wh_-adverbs are generally assumed to undergo LF-movement (or feature movement (Pesetsky 2000, Cheng & Rooryck 2002, Tsai 2004, Soh 2005)) while the _wh_-arguments are not (see also Aoun & Li 1993, Reinhart 1998, cf. Soh 2005), and _wh_-intervention effects are claimed to be a diagnostics on LF-movement (Beck 1996, Beck and Kim 1997) or feature movement (Pesetsky 2000, Cheng & Rooryck 2002, Tsai 2004, Soh 2005). Given this, the ungrammaticality in (6) (Type-II) is unexpected and should be attributed to some other independent factor which is obviously not movement-oriented.

In other words, two types of intervention effects should be categorized. The first one is related to LF-/feature-movement (Type I) while the second one (Type II) has something to do with focus. This amounts to saying that the observations from Beck (1996) and Beck & Kim (1997), and Kim (2002, 2005) and Beck (2006) are basically on the right track. Yet, I will show that the mechanisms behind these effects are different from what they have proposed. More specifically, for the first type, I will adopt the insight of feature movement (Pestsky 2000). I depart from Pesetsky (2000) by suggesting that the Type-I effect should not be triggered by the separation (Pestsky 2000) but by the blocking of feature-movement. As for the Type-II effect, I will show that what is at issue here is the "competition" for one single slot.

3. Minimality Effect (Type I)

As illustrated in (7) the Minimality Effect occurs when the feature movement of an interpretable feature $[iF_1]$ of Y is blocked by an intervening Z bearing a feature of the same sort, $[iF_1]$ (see also Rizzi 2004).

(7) Minimality Effect

   *$[[iF_1]\ X_{[\text{uF1}]} \ldots [Z_{[iF1]} \ldots [\ldots Y_{[\_\_,\ iF2,\ iF3]} \ldots\ ]$

This approach immediately accounts for the contrast between the _wh_-adverb and the _wh_-argument as in (5), assuming Tsai's (1994) parametric approach on _wh_-in-situ construal (see also Reinhart 1998, Pesetsky 2000, Cheng & Rooryck 2002). That is, for the _wh_-argument Chinese employs a base-generated Q-operator merged at C and it unselectively binds the in-situ _wh_-argument to form an operator-variable pair. In (5b) since no movement occurs to the _wh_-argument

Barry C.-Y. Yang

at any level, no intervention occurs. Such an approach fares better than Pesetsky's (2000) separation approach which will in principle rule out the *wh*-argument in (5b).[1]

As for the contrast in (3), I still follow Pesetsky's (2000) observation that the non-*wh*₁-in-situ in English undergoes covert phrasal movement which in the sense of Collins (2004) and Guerzoni (2006) is able to "wrap" the relevant feature within the whole feature bundle package (the whole *wh*-phrase) to get rid of the Minimality Effect. It follows that covert phrasal movement is immune from intervention effects whereas the feature movement is not.

4. Competition Effect (Type II)

The second type of intervention effects involves focus as (6) shows (see also Kim 2002, 2005, Beck 2006). I propose that such a type should result from the Competition Effect which is derived from a traditional notion "one-slot-per-Comp". As illustrated in (8), the Competition Effect occurs when a focus-operator (F-Op) introduced by the focus element and a Q-operator (Q-Op) introduced by the in-situ *wh*-item compete for the same slot, Comp, in CP.

(8) $*[_{CP}$ ___ $[_{IP}$ F-subject$_i$ … *wh*-object$_j$]]
        with F-Op$_i$ and Q-Op$_j$

Though not new, this simple idea elegantly explains the embedding problem of the focus effects that Kim (2002, 2005) and Beck (2006) face (see (4) above) (Tomioka 2007). Here is how it works. When embedded, the F-Op takes the embedded Comp position so that it does not compete with the Q-Op taking the matrix Comp.

(9) $[_{CP}$ Q-Op$_j$ $[_{IP}$ …$[_{CP}$ F-Op$_i$ $[_{IP}$ F-subject$_i$ … *wh*-object$_j$]]

(9) also fits in nicely with Miyagawa & Endo' (2004) observation that the weakening effect in embedded context in Japanese/Korean is in fact due to the D-linking effect. Note also that in the literature the variant speaker judgment is argued to be due to the D-linking (or Specificity) effect (Lee 2001, Kuno & S.-Y. Kim 2004, Miyagawa & Endo 2004, among others). Now, following Pesetsky (1987) the D-linked *wh*-construal is substantiated via a base-generated Q-Op

---

[1] Soh (2005) suggests that Chinese *wh*-arguments should undergo covert phrasal movement following Pesetsky's (2000) reasoning, hence no intervention effect. However, in Yang (2008) I show with ample evidence that Chinese *wh*-arguments should not undergo covert phrasal movement. That's why I abandon the idea of separation and resort to the blocking of (feature-) movement path.

directly merged to CP (or specifically, TopicP) and binds the in-situ *wh*-item. Such a construal is totally in line with (9). That's why D-linking is observed to play a role here. On the other hand, when non-D-linking is forced, the Q-Op in Japanese/Korean type of languages is merged to the in-situ *wh*-item and it has to undergo feature movement to matrix CP (Watanabe 1992, Tsai 1994), crossing the F-Op and triggering the intervention effect.

5. Concluding and Further Remarks

This study shows that, firstly, categorizing two types of intervention effects helps sort out the puzzling variations both across and within languages. Secondly, each type is modeled with independent triggering factors which better capture the language paradigm than currently available literature.

One last thing to note: we will assume the D-linking construal in Japanese/Korean type of languages differs from that in English type since Pesetsky (2000) shows convincingly the D-linked *wh*-in-situ in English is subject to either feature movement of covert phrasal movement. In this sense, (10a) falls out naturally with either the Minimality or Competition Effect, assuming the feature (or Q-Op) movement of *which girl* (Pesetsky 2000). On the other hand (10b) is immune from Competition Effect because assuming *which boy* undergoes covert phrasal movement (Pesetsky 2000) such a movement targets the Spec of CP whereas the F-Op targets the Head/Comp of CP, hence no competition.

(10) a. ??Which boy did <u>only Mary</u> introduce **which girl** to __.
     b.   Which girl did <u>only Mary</u> introduce ___ to **which boy**?

References

Aoun, Joseph, and Audrey Li. 1993. Wh-elements in situ: Syntax or LF. *Linguistic Inquiry* 24: 199-238.

Beck, S. and S.-S. Kim. 1997. On wh- and operator scope in Korean. *Journal of East Asian Linguistics* 6:4, 339-384.

Beck, Sigrid. 1996. *Wh-constructions and transparent Logical Form*. Ph.D. dissertation, University of Tuebingen.

Beck, Sigrid. 2006. 'Intervention effects follow from focus interpretation'. Natural Language Semantics 14, 1-56.

Cheng. Lisa. 1991. *On The Typology of wh-questions*, Cambridge, MA: MIT Dissertation.

Cheng, Lisa. and J. Rooryck. 2002. Types of *wh*-in-situ, ms. Leiden University.

Collins, Chris. 2004. A *smuggling* approach to raising in English. *Linguistic Inquiry* 36.2, 289-298. 2004

Guerzoni, Elena. 2006. Intervention effects on NPIs and feature movement: Towards a unified account of intervention. *Natural Language Semantics*, 14(4):359-398.

Kim, S.-S. 2002. Intervention effects are focus effects. *Japanese/Korean Linguistics* 10, CSLI Publication, 615-628.

Kim, S.-S. 2005. Focus intervention effects in questions. Paper presented at TEAL Workshop 3, Harvard University.

Kuno, Susumu and Soo-Yeon Kim 2004. On syntactic intervention. In: Susumu Kuno et al. (ed.), *Harvard Studies in Korean Linguistics*, Vol. 10. Harvard University: Dept. of Linguistics, pp. 3–31.

Lee, Mina. 2001. Intervention Effects Revisited. ed. by Susumu Kuno *et. al*. *Harvard Studies in Korean Linguistics* 9:563-572.

Miyagawa, Shigeru. 2004. Intervention Effects Are Syntactic, class handout.

Miyagawa, Sigeru, and Yoshio Endo. 2004. Intervention effects are not pragmatic, ms. MIT and Yokohama National University.

Pesetsky, David. 1987. Wh-in-situ: movement and unselective binding. In Eric J. Reuland and Alice GB ter Meulen (eds.), *The Representation of (In)definiteness*, pp 98-129. Cambridge, MA: MIT Press.

Pesetsky, David. 2000. *Phrasal Movement and Its Kin*, MIT Press.

Reinhart, Tanya. 1998. Wh-in-situ in the framework of the minimalist program. *Natural Language Semantic*s 6:1, 29-56.

Rizzi, Luigi. 2004. On the form of chains: Criterial positions and ECP effects, ms.

Soh, Hui-Ling. 2005. WH-in-situ in Mandarin Chinese. *Linguistic Inquiry* 36: 143-155.

Tomioka, Satoshi. 2006. LF intervention effects: Old and new challenges, paper presented at Workshop in Current Issues in Semantics.

Tomioka, Satoshi. 2007. Pragmatics of LF intervention effects: Japanese and Korean interrogatives, *Journal of Pragmatics,* 39: 1570-1590.

Tsai, Wei-Tien Dylan. 1994. *On Economizing the Theory of A-bar Dependencies*. Doctoral dissertation, MIT, Cambridge, Mass.

Watanabe, Akira. 1992. Subjacency and S-structure movement of wh-in-situ. *Journal of East Asian Linguistics* 1, 255-291.

Yang, Barry C.-Y. 2008. *Intervention Effects and the Covert Component of Grammar*. Doctoral dissertation, National Tsing Hua University, Hsinchu, Taiwan.

Sentential Adverbs as Focusing Adverbs

Chih-hsiang Shu
Stony Brook University, New York

Abstract

In this paper I argue for the view that cross-linguistically, sentential adverbs are inherently focus-sensitive adverbs. Evidence for this view comes from both semantic and syntactic considerations cross-linguistically. Specifically, it will be clear that S-adverbs have the distributional freedom and constraints of typical focusing adverbs. It is also argued that current theories of left-periphery, including Agree (Chomsky 2000, 2001, Pesetsky and Torrego 2007) and adjunction (Chomsky 2004) offer us some simple ways of accounting for the facts which cannot be easily accounted for by previous approaches such as the Split-IP analysis (Cinque 1999, inter alia), covert movement analysis (Shu 2006), and overt movement plus remnant movement analysis (Kayne 1998).

1. Introduction

The main goal of this study is to account for a family of puzzles of sentential adverbs (S-adverbs) that have not been properly accounted for in the literature. First, sentential adverbs (S-adverbs) seem to occur in the 'wrong' position in many cases, as illustrated in the following:

(1) a. John *certainly* saw Mary.
  b. Bill offended *possibly* everyone.
  c. John and *maybe* Mary went to the store.   (Collins 1988)
  d. You *either* leave now or I'll call the police.   (Den Dikken 2006)

These positions are not the 'right' positions because as propositional operators, S-adverbs are supposed to only occur in the sentence-initial position, just like a predicate taking a DP argument should occur outside the DP, but not within the DP. However, in (1), we see that S-adverbs can not only occur after the subject (1a, d), but also occur in a more deeply embedded position (1b, c). Recent syntactic analyses, such as the split-IP functional hierarchy analysis (Cinque 1999, et al.), or the semantically-oriented analysis (Ernst 2002) fail to properly account for such facts, if at all.

Chih-hsiang Shu

Second, it is also not explained why cross-linguistically, S-adverbs are *always* able to occur between the verb phrase and the subject (again the 'wrong' position), but *not always* able to occur in the pre-subject position:

(2)   a. *Who did <u>possibly</u> John see?
      b. Who did John <u>possibly</u> see?   (cf. Svenonius 2002 and Den Dikken 2006)

(3)   a. zhangsan <u>yiding</u>   kandao-le lisi   (Chinese)
         Z.   certainly see-Asp   L.
         'Zhangsan certainly saw Lisi.'
      b. *<u>yiding</u> zhangsan   kandao-le lisi

(4)   a. Núna hafa   bófarnir   líklega   stolið smjörinu   (Icelandic)
         now   have   the.gansters probably   stolen the.butter
         'Now probably the gangsters have stolen the butter.'
      b. ?? Núna hafa <u>líklega</u>   bófarnir   stolið smjörinu
            now have probably the.gangsters stolen the.butter (Svenonius 2002)

The question is: why is the 'unmarked' position of S-adverbs the post-subject position?

Another puzzle is that while S-adverbs can be deeply embedded in a sentence in languages like English, as shown in (1), they cannot be so deeply embedded in languages like Chinese:

(5)   a. zhangsan (xiangran) hui (*xiangran) qu   meiguo   (Chinese)
         Z.   obviously will obviously go.to   U.S.
         'Zhangsan will obviously go to the U.S.'
      b. zhangsan (yiding)   renzhi (*yiding) genduo ren
         Z.   certainly know   certainly more   person
         'Zhangsan certainly knows more people.'

The question is: what parameters condition the cross-linguistic differences?

In this paper I will argue for a less well-know view of S-adverbs, namely S-adverbs are inherently focusing adverbs (Engels 2005, Shu 2006). In section 2, I will review the fundamental properties of S-adverbs as are known in the literature, and show how the literature have conflicting views based on different sets of data. In the next 2 sections, I focus on the focus-sensitivity property in more detail. In section 3, I discuss the current understanding of focus and association with focus. In section 4, I provide arguments for the view that S-adverbs are focus-sensitive adverbs. In section 5 and 6, I lay down my proposal to solve the problems mentioned above, based on recent developments of Agree

(Chomsky 2000, 2001, Pesetsky and Torrego 2007) and adjunction (Chomsky 2004) in the minimalist program. Section 7 concludes the paper.

2. Some basic facts about sentential adverbs

In this section I will review some fundamental syntactic and semantic properties of S-adverbs that prove to be problematic for a developing a coherent theory so far.

2.1 S-adverbs scope over the whole proposition

S-adverbs convey the speaker's attitude, evaluation, or some other modes of perception about the whole proposition. Syntactically, they behave differently from adverbs that don't take propositional scope.

(6) a. Did they run out of fuel quickly?
     b. *Did they probably run out of fuel?   (Nuyts 1993)

(7) a. They didn't completely run out of fuel.
     b. *They didn't certainly run out of fuel.

(8) a. They solved the problem not only quickly but also simply.
     b. *They have not only possibly but even probably run out of fuel.
        (ibid.)

The above contrasts suggest that since S-adverbs take a wide scope, since they cannot appear certain in syntactic positions where 'narrow-scope' adverbs can occur.

2.2 S-adverbs precedes other classes of adverbs

A relevant fact suggesting S-adverbs take propositional scope is that they have to precede non-S-adverbs when both types of adverbs occur in the same sentence, shown as follows:

(9) a. John has perhaps already finished cooking.
     b. * John has already perhaps finished cooking.

(10) a. John will certainly often see Mary.
     b. *John will often certainly see Mary.

Based on this 'rigid ordering', some linguists (Cinque 1999 et al.) argue that the relevant adverbs are specifiers of functional heads, which are ordered with respect to each other by a universal functional hierarchy. This split-IP analysis, however, wrongly predicts the default position of S-adverbs is the sentence-

initial position (since the functional heads checking mood features is higher than T_0, the specifier of which is the subject by default). In the next section I will show how this is not true.

2.3 Sentence-initial position is not the 'unmarked' position

As we have seen in the introduction, S-adverbs do not always occur in the sentence-initial position, and they are sometimes actually bared in that position. As shown in (1), they can be deeply embedded in languages like English. In (2-4), we see that they are sometimes barred in the sentence-initial position. It is clear that both facts pose problems for the universal functional hierarchy analysis. For example, it is not clear how an element bearing propositional scope should occur low in the sentence structure. Even if we allow movements of subjects, auxiliaries, and sometimes even verbs to positions before the S-adverbs, it is not clear how the movements are obligatory sometimes.

2.4 S-adverbs are focus-sensitive

It has been noted by several German linguists (Lang 1979, Jacobs 1983, 1986, Engels 2005) that S-adverbs behave like focus-sensitive adverbs. For example, Engels (ibid.) has the following observation, with examples given in (12):

(11) [Sentence adverbs] tend to occur left-adjacent to their associated
       constituent, following unfocused material.

(12) a. (What happened yesterday?)
        Gestern hat (vermutlich) Karl (??vermutlich) Fritz (??vermutlich) ein Auto
        Yesterday has presumably Karl   Fritz   a car
       (*vermutlich) geschenkt given
      b. (What did Karl do yesterday?)
        Gestern hat (*vermutlich) Karl (vermutlich) Fritz (??vermutlich) ein Auto
         (*vermutlich) geschenkt
      c. (What did Karl give to Fritz yesterday?)
        Gestern hat (*vermutlich) Karl (*vermutlich) Fritz (vermutlich) ein Auto
         (*vermutlich) geschenkt

S-adverbs are focus-sensitive in (11) in the sense that their syntactic position is sensitive to the locus of the information focus in the sentence. Similarly, Shu (2006) also observes that focus-sensitivity of S-adverbs is manifested in Chinese:

(13) a. In Chinese, some S-adverbs are allowed to be left-adjacent to the subject
        only if the latter is marked by focus.

    b. S-adverbs are left-adjacent to information focus in answers to wh-questions. (cf. also Svenonius 2002)

    c. S-adverbs follow TP-internal topics in Chinese.

    d. S-adverbs of the main clause follow *when-*, *after-*, and *before-* clauses unless the former are topicailzed.

    e. S-adverbs precede focus-markers such as *shi* in Chinese.

In view of these properties, both Engels and Shu propose that S-adverbs are base-merged at the edge of *v*P or TP, depending on which XP is the locus of information focus. Shu further proposes that S-adverbs later move covertly to the edge of CP to check the [+mood] feature. Their proposals have the merit of covering a wider range of data than the split-IP alternative. It is not clear, however, how the rigid ordering of adverbs should be properly explained, and why we can't find languages where movements have to be overt, and what exactly 'focus-licensing' means.

      To sum up this section, we have seen that S-adverbs have a mixture of properties that have been accounted for by a mixture of theories. On the one hand, their wide semantic scope and high position with respect to other classes of adverbs support a split-IP analysis where S-adverbs occupy fixed spec-of-IP positions. On the other hand, their focus-sensitivity suggests an analysis where S-adverbs have various possible IP or non-IP base positions. These conflicting theories clearly need to be replaced by a unified theory, which is the goal of this paper. Before doing so, I will provide some more arguments to show S-adverbs are indeed focus-sensitive in the next two sections.

3. The defining properties of focus and association with focus

What is focus? And what is focus-sensitivity? In this section, we will look at each notion carefully, and then examine some common semantic and syntactic properties that are related to focus and focus-sensitivity.

3.1 The basic definition

Krifka (2007) has a very clear definition of focus:

(14) Focus indicates the presence of alternatives that are relevant for the interpretation of linguistic expressions.

In addition, Krifka defines focus-sensitive adverbs as follows:

(15) If the interpretational effect of an adverb depends on the (information or contrastive) focus, the expression is a focus-sensitive adverb.

A typical example of focus-sensitive adverbs is *only*. Its interpretational effect clearly depends on the focus of the sentence, as illustrated below:

(16) a. John only READS novels. (He doesn't write them.)
   b. John reads only NOVELS. (He doesn't read newspapers.)

Beyond *only*, many kinds of expressions have been identified as focus-sensitive in the literature. They include the generic operator and adverbial quantifiers (Rooth 1985, Krifka 1995), illocutionary operators (Jacobs 1984), a variety of propositional operators, including negation, modals, sentence adverbs, and sentence-embedding verbs (Jackendoff 1972, König 1991, Rooth 1996, Brennan 1997), a variety of adverbs like *merely*, *truly*, *simply*, and *hardly* (Jackendoff 1972), reason clauses (Krifka 2007), and adjectives like *very* and *mere*.

3.2 Syntactic idiosyncrasies

In addition to the defining properties above, typical focus-sensitive adverbs also have some idiosyncratic properties. Linguists have yet to account for these properties in a systematic way, but some descriptive generations prove to be useful for our purpose:

(17) a. Adverbs like *only* and *even* can attach to various syntactic categories, but not to IP in English (Bayer 1996), and not to post-verbal nominal objects in Chinese. *Either*, on the other hand, can attach to IP (Den Dikken 2006).
   b. A focusing adverb must c-command a focused constituent (Tancredi 1990, Bayer 1996), unless the former is in the auxiliary complex and the latter is the subject (this caveat doesn't apply to the adverb *only*[1]) (cf. König 1991).

To illustrate, attachment to DP is shown below:

(18) a. John read [only NOVELS]. (=16b)
   b. Did [only JOHN] see Mary?
   c. What did [either JOHN] see or MARY?   (Den Dikken 2006)

Attachments to DP objects in Chinese, on the other hand, are barred.

(19) a. zhangsan zhi   kandao LISI
      Z.   only see   L.
     'Zhang saw only LISI.'

---

[1] In Chinese this caveat also doesn't apply to *shenzhi* 'even'. It has to c-command the subject if the latter is the focused associate.

   b. *zhangsan kandao zhi(-you)[2] LISI
     Z.   see   only-Prt   L.

Attachments to *v*P, T′, TP, and CP are illustrated below:

(20) a. John [only saw MARY].   (*v*P)
     b. John [only can see Mary].   (T′)
     c. [Either John ate rice or beans].   (TP)
     d. [Even if you are right], you can't shout. (CP)

The lexically-conditioned ban on TP adjunction is illustrated as follows:

(21) a. *Even John saw MARY.
     b. *Only John saw MARY. (ok with non-exhausitive-identification meaning)

The c-commanding requirement is illustrated below:

(22) a. *[JOHN] likes even Mary.
     b. *John [READS] only novels.
     c. [JOHN] can even write novels.
     d. *John admires Susan now, and he [HAS] always admired her.

                                          (Baker 1971)

3.3 Non-focusing adverbs lack these properties

Any analysis will not be complete without a control group. Can the above defining properties of focus-sensitivity be applied to non-focus-sensitive adverbs? When we test them on manner adverbs and temporal adverbs, we see the answer is clearly no. First, their interpretational effect is not dependent on the focus.

(23) a. JOHN read the novel quickly.
     b. John READ the novel quickly.

In these two sentences, the focus of the sentence can only associate with the assertion operator and express the emphasis of the speaker's assertion, but not associate with *quickly*.

    Second, those adverbs also do not have the freedom of attaching to different syntactic categories as focusing adverbs do.

(24) a. *John read [quickly the NOVEL.]
     b. *Did [quickly JOHN] read the novel?
     c. John (*carefully) can (*carefully) work.

---

[2] When *zhi* precede a DP, it has to be followed by the particle *you* or *shi*.

Chih-hsiang Shu

Thus it seems clear that focus-sensitivity is a valid property that helps us distinguish between different types of adverbs. In the next section, we will see that S-adverbs are focus-sensitive by these definitions.

4. S-adverbs are focus-sensitive adverbs

It should be now clear how focus and focus-sensitivity are manifested in the semantics and syntax of certain expressions. Now let's see whether S-adverbs have the relevant properties.

4.1 The interpretational effect of S-adverbs

Krifka (2007) argues that S-adverbs are focus-sensitive based on the semantic criterion (15). His arguments come from the following example:

(25) Fortunately, Bill spilled WHITE wine on the carpet.

According to Krifka, a proper understanding of (25) is as follows: among two alternatives, BILL SPILLED RED WINE BILL and BILL SPILLED WHITE WINE, the latter one was more fortunate. Although this is a very rough sketch of semantics, it does seem to capture an important inherent property of S-adverbs that has been ignored in the bulk of literature.

4.2 Attachment to various syntactic categories and some lexical ideosyncracies

Syntactic facts also suggest that S-adverbs are focus-sensitive. Like typical focusing adverbs, S-adverbs can attach to DP, vP, T′, and TP:

(26) a. John and [maybe Mary] went to the store.   (=1c) (DP)
    b. Who did [possibly JOHN] see?   (DP)
    c. John [certainly saw Mary].   (vP)
    d. You [either leave now] or I'll call the police. (=1d) (vP)
    e. John [certainly could see Mary].   (T′)
    f. [Certainly John could see Mary].   (=6a) (TP)

Just like typical focusing adverbs in (19), S-adverbs are also banned to occur in a post-verbal object position in Chinese:

(27) a. zhangsan dagai   jinu-le meiyige ren
      Z.   possibly irritate every   person
     'Zhangsan possibly irritates everyone.'
   b. *Zhangsan jinu-le dagai meiyige ren

The lexically-conditioned ban on TP-adjunction applies to some S-adverbs in Chinese (but not generally in English[3], presumably due to the obscuring lexically encoded topicalization possibilities.

(28) a. *[yiding   zhangsan kandao-le lisi]   (=3b)   (TP)
          certainly Zhangsan see-Asp   Lisi
     b. zhangsan [yiding   kandao-le lisi]   (=3a)   (vP)
        Zhangsan certainly see-Asp   Lisi
        'Zhangsan certainly saw Lisi.'

That (28a) is ungrammatical because *yiding* cannot be topicalized is witnessed by its inability to take the topicalization suffix –a, in contrast with other *topicalizable* S-adverbs.

(29) a. *yiding   (a),   zhangsan kandao-le lisi
          Certainly Top Zhangsan see-Asp   Lisi
     b. xianran (a),   zhangsan kandao-le lisi
        obviously Top Zhangsan see-Asp   Lisi
        'Obviously, Zhangsan saw Lisi.'

Finally, like typical focusing adverbs, S-adverbs must c-command their focused associates, unless they are in the auxiliary complex.

(30) a. *[Everyone] likes obviously Mary.
      b. *John [READS] certainly novels.
      c. [JOHN] can definitely write novels.
      d. *John DID obviously see Mary.

Examples in (12) and (13) show the same pattern.

To summarize, in this section we see S-adverbs does conform to the defining properties of focusing adverbs enumerated in section 3.

Now, with solid evidence that S-adverbs are focusing adverbs, we are in a position to offer a proper theoretic analysis for S-adverbs and focusing adverbs in general.

5. An account under the theory of Agree

Let's recap what we have learned about S-adverbs so far:

---

[3] *Sure as hell* is one of the S-adverbs that can't adjoin to TP in English, as is illustrated below:
(i) a. John sure as hell likes Mary.
    b. *Sure as hell, John likes Mary.

(31) (i) S-adverbs scopes over the whole proposition.
      (ii) S-adverbs always precede other classes of adverbs with narrower scopes.
      (iii) S-adverbs can occur in a low position with respect to non-adverbial elements in a sentence.
      (iv) S-adverbs, and focusing adverbs in general, are subject to other different syntactic constraints in different languages (e.g. the attachment site in Chinese is more restricted).
      (v) The c-commanding requirement (17b) is not always observed.

Equipped with a better understanding of the distributional properties of S-adverbs, we are now in a position to offer a proper theoretical analysis. First, let's review briefly some recent developments of the minimalist program that has some bearing on our analysis.

5.1 Current theoretical understanding of the syntax of adjunction, the syntax of focusing adverbs, and the syntax of focus-inversion

5.1.1 Adjunction: Properties and Theories

Although it has been the goal of certain syntactic tradition of eliminate adjunction as a theoretical entity in the generative grammar (Kayne 1994 et al.), the core properties of adjunction have not yet been successfully reduced to independently well-motivated theories of pure Merge. The following two properties are a case in point:

(32) a. After merging with an XP, adjuncts do not project.
     b. Generation of adjuncts during the derivation is counter-cyclic. (Lebeaux 1988, Chomsky 2004)

S-adverbs and focusing adverbs clearly do not project and can be inserted late in the derivation (hence their syntactic 'low' position), so they should be treated as adjuncts instead of specifiers or functional heads.

A recent theory of adjunction is given in Chomsky (2004). According to him, in a nominal [DET <ADJ, NP>], ADJ is adjoined to NP, which all then undergoes set-Merge with determiner DET. We thus have the structure {DET, <ADJ, NP>}, in which NP retains the properties it would normally possess in non-adjoined structures, and ADJ can be seen as occupying a 'separate plan'. In addition, he also talks a little bit about the counter-cyclicity property. Although no formal accounts are given, he makes connections with respect to apparent rightward adjunct extraposition and the ACD constructions.

5.1.2 Agree: The Theory and its Applications

We have noted the tension between the *high* semantic scope and the *low* syntactic position. It has also been independently noted the syntax of adjunction allows counter-cyclicity. Do we need to create an *ad hoc* syntactic mechanism to deal with the facts? Fortunately, we don't, since there is already an independently well-motivated theory that allows exactly this kind of situation. The theory is called *Agree* (Chomsky 2000, 2001, Pesetsky and Torrego 2007), which deals with general *dependency* relationships between two distinct syntactic elements. To specifically, Chomsky defines Agree as follows (ibid.):

(33) a. Agree establishes a relation (agreement, Case checking) between an LI α and a feature F in some restricted search space.
      b. Agree involves the erasure of uninterpretable features of probe and goal.

In the same vein, Pesetsky and Torrego (ibid.) defines a more detailed feature sharing version of Agree:

(34)   a. An unvalued feature F (a *probe*) on a head H at syntactic location α (F_α) scans its c-command domain for another instance of F (a *goal*) at location β (F_β) with which to agree.
       b. Replace F_α with F_β, so that the same feature is present in both locations.

It should be clear now that the so-far-mysterious placement of S-adverbs and focusing adverbs in general doesn't seem to be so mysterious anymore under the theory of Agree. In fact, analyses in terms of Agree have been proposed about focusing adverbs (particles) like *no* and *only* in some recent studies (Watanabe 2004, Horvath 2007). Horvath's analysis of *only* in Hungarian, in fact, is very close to the analysis I'll propose for S-adverbs in the next section. Let me cite the following passage as a 'sneak preview' of my analysis below (ibid.: 129):

(35) a. Assume that there is an exhaustive identification (EI) operator, and a clausal functional head EI^0 with an uninterpretable EI-operator feature. This operator feature of EI_0 enters into a matching ('Agree') relation with a phrase in its search (c-command) domain. The EI^0 head has furthermore an EPP feature, and consequently movement, rather than merely Agree applies: a matching EI-Op phrase gets attracted by the EI^0 to the Spec, EIP position.
     b. The EI operator (able to enter into an Agree relation with the corresponding [EI] feature of the clausal EI^0) can be merged into structure at the root of DP (and apparently of some other phrasal categories, such

as PP, VP, and CP, as well). Thus EI-OP movement pied-pipe the phrase whose outermost specifier it occurs in.

c. The EI-Op requires the presence of stress-based, "information" focus within its c-command domain, namely within the phrase it attaches to, just like ONLY and EVEN do within theirs, i.e. it manifests the property of association with focus.

5.2 An agree analysis of S-adverbs

Based on the recent developments of the minimalist program and (31), we now present our proposal:

Revised feature sharing version of Agree

(36) a. A feature F (a *probe*) on a head H at syntactic location α (F_α) scans its c-command domain for another instance of F (a *goal*) at location β (F_β) with which to agree.

b. Replace any unvalued feature with valued feature, so that the same feature is present on both α and β.

An Agree analysis of S-adverbs

(37) a. (i) C   D/V/Aux, etc. $\rightarrow$ C   D/V/Aux, etc.
[valued *i*Mood] [unvalued *u*Mood] [valued *i*Mood] [valued *u*Mood]
(ii) Condition: A *goal* has to bear focus, unless it is the Aux or the main verb.

b. This Agree operation then triggers pair-Merge, realized as late insertion of S-adverbs at the edge of the *goal*. (e.g. {<S-ADV, V>, DP})

c. There is a fixed universal functional hierarchy than governs the order of C_0 heads. The rigid ordering of focus-sensitive adverbs is a reflex of this fixed hierarchy.

The purpose of the revision in (36) is to give a less restricted definition of probe and goal than P&T's version, so the *goal* can be unvalued before Agree applies. The reason is it's not clear how the adjunction sites of S-adverbs are inherently valued, rather than the C_0 itself. In (37), S-adverbs are basically treated as the reflex of Agree, just like inflectional and tense suffixes marked on verbs or auxiliaries in English.[4] C_0 heads are ordered with respect to a fixed hierarchy a la Rizzi (1997) and Haegeman (2000a,b).

---

[4] The difference is that they can attach not only to the first verbal element, but also the post-verbal elements, and also phrasal categories. In addition, they can also be topicalized, as we have

It is clear that all the basic facts and facts about focus-sensitivity in section 2 and 3 can be accounted for by (37). First, the valued interpretable feature of [mood] is located at C_0, so it follows naturally that S-adverbs take propositional scope. Second, the C_0 head triggering later insertion of S-adverbs are higher than the functional heads that are relevant to adverbs taking narrow scopes, the ordering effects follow naturally. Third, since the overt syntactic positions of S-adverbs are the reflexes of Agree, not first Merge, the sentence-initial position is not an unmarked position. Fourth, the focus-sensitivity, although not fully explained yet, can be subsumed under the independently well-motivated Agree theory, just like the analysis offered by Horvath for typical focusing adverbs in (35). Fifth, the cross-linguistic variation as witnessed in (5) can now be related to the fact Chinese lexical categories are relatively 'virus-free' (Huang 2005), unlike languages like English. For this reason, not as many lexical categories in Chinese bear unvalued uninterpretable features as in English. Therefore, both S-adverbs and focusing adverbs cannot attach 'low' in Chinese.

6. Some novel predictions and consequences

This analysis, if correct, has some consequences that can help us understand some other properties of S-adverbs and focusing adverbs in a coherent picture, and the theory of Agree in general. First, an Agree analysis predicts the possibility of Move after S-adverbs are merged. That is, in the configuration

(38) $[_{CP}C_0\ldots[\text{S-adv-XP}]]$

Where C_0 agrees with XP, we expect XP to be able to undergo Move in certain languages, perhaps along with the attached S-adverb.

(39) $[_{CP}[\text{S-adv-XP}]_i C_0\ldots t_i]$

This state of affairs is borne out in English and Cantonese.

(40) a. [Not until yesterday]$_i$ did he arrive t_i.
    b. [mou fan aa]$_i$ keoi gingjin t_i!
      Neg eat Prt he   surprisingly
      'Surprisingly, he didn't eat.'
    c. [sausi]$_i$ lo1 keoi m   zungji zing t_i.   (Law 2003)
      sushi Prt he Neg like   make
      'S/he doesn't like to make SUSHI (as opposed to dumplings).'

---

seen above. These differences need a proper theoretical account, which I leave for future research.

Although the S-adverbs in (40b) and the focusing adverb in (40c) do not move along with the fronted *goal*, it is clear that not only Agree, but also Move may be involved in the syntax of S-adverbs.

Second, the Agree analysis also predicts the existence of *focus concord*, where two elements in the same sentence are marked for focus, but there is only one semantic focus involved. This prediction is again borne out in Chinese:

(41) zhiyou ZHANGSAN cai   neng   fu   ci   zhongren
     only   Z.   CAI able.to bear this grave.responsibility
     'Only Zhangsan can bear this grave responsibility.'

The concord-like multiple-Agree is realized both on the subject *Zhangsan* and the auxiliary *neng* in (41). Interested readers can refer to Hole (2004) for a detailed discussion, who also adopts an Agree analysis for particles like *cai*.

Third, from a macroscopic perspective, it is only reasonable that we apply similar analyses to various other expressions that show tension between 'high' and 'low' positions. These may include:

(42) a. Mood (or other elements with a high scope) affixes/clitics on DP
       arguments. (Kayardild 'mood' case marking, Icelandic case marking
       conveying aspectual information, Spanish prepositions conveying
       aspectual information, English negative marker *no*, prefix *wh-*, Chinese
       focus particles *shi/you/hui*, Korean 'extrinsic' plural suffix *-tul*)
   b. 'High scope' adjectives on NP's. (*John takes an occasional shower.*)
   c. Mood or high scope clitics on verbs. (Chinese 'potential' construction V-
      de/*bu*-A, mood clitic *ge*, English *seem to* (cf. *Mary still can't seem to play
      soccer*)).

If all the above morphosyntactic elements can be accommodated under the Agree theory, than we can have a really 'minimalist' theory that captures the 'maximal' coverage of empirical facts.

7. Conclusion

In this paper I have shown that there are strong pieces of evidence suggesting S-adverbs are inherently focus-sensitive adverbs. In addition, I have suggested the dependency relation between the 'high' and the 'low' positions of S-adverbs can best be accommodated by the theory of Agree, a well-established and independently-motivated theory, and the theory of pair-Merge, a promising but yet underexplored territory. I have also suggested that a universal functional hierarchy analysis a la Cinque (1999) should be incorporated into our analysis to explain the rigidity effect. In addition, I have suggested the independently-

motivated analytic/synthetic parameter also plays a role in the cross-linguistic variations of the attaching sites of S-adverbs and focusing adverbs. Thus, hopefully, the puzzling empirical and theoretical issues of S-adverbs and focusing adverbs mentioned in the introduction are no longer beyond a proper theoretic account.

The achievements aside, there are, of course, many outstanding issues. What are the proper semantic representations of the various S-adverbs? Are they truth-functional? Why does Agree involve focus-sensitivity? When do we have EPP and multiple-Agree? What more morphosyntactic properties of S-adverbs are out there? I leave them for future research.

Appendix: The Agree approach vs. the alternative approaches

In this section let's look at some alternative analyses of S-adverbs and focusing adverbs that have been proposed in the literature and see why none of them are able to deal with all the empirical and theoretical issues mentioned in this study. The alternatives I will discuss include the following:

(43) a. Covert movement of the S-adverbs (Shu 2006).
     b. Overt movement of the S-adverbs and adjacent elements, followed by remnant movement of elements of the rest of the sentence (a la Kayne's (1998) analysis for focusing adverbs)
     c. There are no syntactic operations at all. Either the apparent syntactic dependencies are in fact operations in the semantic component, or S-adverbs actually take a low semantic scope (a la Ladusaw's (1988) analysis of quantificational adverbs and Laka's (1990: 82) analysis of negation).

I will briefly discuss the problems with these analyses.

A1 Covert movement approach

As I have briefly discussed in 2.4, the covert movement analysis is based on the fact that S-adverbs are interpreted at a 'high' position, but realized at a 'low' position. In this approach, a focusing adverb is adjoined to a FocP in overt syntax, and move to spec-of-CP covert to check the [mood] feature. The problems, in addition to the ones in mentioned in 2.4, are as follows:

(44) a. First Merge is typically θ-related. Adjoining an S-adverb to a FocP is clearly not.

b. The first Merge of S-adverbs are conditioned by something merged later in the derivation (C), this violates cyclicity.

c. Attaching an S-adverb to an object DP is banned in certain languages. This cross-linguistic variation is not expected.

A2 Overt movement plus remnant movement approach

According to Kayne (1998), a focusing adverb in a sentence *John likes only beer* undergoes first Merge with a focused XP, then they move togother to the specifier position of a higher syntactic head. This is then followed by remnant movement of elements if the rest of the sentence. The problems of this approach are as follows:

(45) a. Both the subject and the verb have to move to the edge of CP. These movements are not well-motivated.

b. The problem of first Merge mentioned in (44) still exists.

c. Cross-linguistic variations are still unaccounted for.[5]

d. It is unclear how focus inversion should be derived.

e. This approach involves more complicated derivations than either the covert movement or the Agree analyses.

A3 No operation approach

According to this approach, S-adverbs are simply base-generated in a position lower than TP. Neither Agree nor covert movement is involved. This approach, although simple and seems *prima facie* attractive, has the following problems:

(46) a. It is simply not explained what motivated the first Merge at all. No theoretical interesting questions are asked.

b. The cross-linguistic variation on the attaching site is unexpected if only pure Merge is involved.

c. That S-adverbs can be overtly realized at C_0 in certain languages, and that focus inversion exists in some languages, are unexpected.

---

[5] Huang (2003) adopts Kayne's analysis and suggests that it implies English allows remnant movement whereas Chinese doesn't, and goes on to connect this difference to the fact that Chinese is more analytic and English is more synthetic. It is not clear to me this is a valid connection, since it has not been established what motivates remnant movement at all.

Sentential Adverbs as Focusing Adverbs

References

Baker, Chris L. 1971. Stress level and auxiliary behavior in English. *Linguistic Inquiry* 2:167-181.

Bayer, Josef. 1996. *Directionality and logical form: On the scope of focusing particles and wh-in-situ*. Dordrecht: Kluwer Academic Publishers.

Brennan, Virginia. 1997. Quantification modals. *Linguistic Inquiry* 28:165-169.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, ed. by Roger Martin, David Michaels, and Juan Uriagereka, 89-155. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A life in language*, ed. by Michael Kenstowicz, 1-52. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 2004. Beyond explanatory adequacy. In *Structure and beyond: The cartography of syntactic structures, volume 3*, ed. by Adriana Belletti, 104-131. Oxford: Oxford University Press.

Cinque, Guglielmo. 1999. *Adverbs and functional heads: A cross-linguistic perspective*. New York: Oxford University Press.

Collins, Christopher. 1988. Conjunction adverbs. Ms., MIT, Cambridge, Mass.

Den Dikken, Marcel. 2006. *Either*-float and the syntax of co-*or*-dination. *Natural Languages & Linguistic Theory* 24: 689-749.

Engels, Eva. 2005. *Adverb placement: An optimality theoretic approach*. PhD dissertation, University of Potsdam.

Ernst, Thomas. 2002. *The Syntax of Adjuncts*. Cambridge: Cambridge University Press.

Haegeman, Liliane. 2000a. Inversion, non-adjacent inversion, and adjuncts in CP. *Transactions of the Philological Society*. 98.1: 121- 160.

Haegeman, Liliane 2000b. Negative preposing, negative inversion, and the Split CP. In *Negation and Polarity*, ed. by Lawrence R. Horn and Yasuhiko Kato, 21-61. Oxford: Oxford University Press.

Horvath, Julia (2007). Separating "focus movement" from focus. In *Phrasal and clausal architecture: Syntactic derivation and interpretation. In honor of Joseph E. Emonds*, ed. by Karimi, Simin, Vida Samiian, and Wendy K. Wilkins, 108-145. Amsterdam: John Benjamins.

Huang, C.-T. James. 2003. The distribution of negative NPs and some typological correlatives. In *Functional structure(s), form and interpretation: Perspectives from East Asian languages*, ed. by Audrie Y.-H. Li and Andrew Simpson, 262-280. London: Routledge.

Huang, C.-T. James. 2005. On syntactic analyticity and the other end of the parameter. Lecture notes from LSA 2005 Linguistic Institute course. Ms., Harvard University.

Jackendoff, Ray S. 1972. *Semantic interpretation in generative grammar*. Cambridge, MA: MIT Press.

Jacobs, Joachim. 1983. *Fokus und Skalen: Zur Syntax und Semantik der Gradpartikeln im Deutschen*. Tübingen: Niemeyer.

Jacobs, Joachim. 1984. The syntax of bound focus in German. *Groninger Arbeiten zue Germanistischen Linguistik* 24, 172-200.

Jacobs, Joachim. 1986. The syntax of focus and adverbials in German. In *Topic, focus, and configurationality*, ed. by Werner Abraham and Sjaak de Meij, 103-127. Amsterdam: Benjamins.

Kayne, Richard. 1998. Overt vs. covert movement. *Syntax* 1:128-191.

König, Ekkehard. 1991. *The meaning of focus particles: a comparative perspective*. London: Routledge.

Krifka, Manfred. 1995. Focus and interpretation of generic sentences. In *The Generic Book* , ed. by Carlson, Gregory N. and Francis Jeffery Pelletier. 238-264. Chicago: The University of Chicago Press.

Krifka, Manfred. 2007. Basic notions of information structure. In *Working Papers of the SFB632, Interdisciplinary Studies on Information Structure* (ISIS) 6, ed. by Caroline Féry, Gisbert Fanselow, and Manfred Krifka, 13-56. Potsdam: Universitätsverlag Potsdam.

Laka, Mugarza. 1990. *Negation in syntax: On the nature of functional categories and projections*. Ph.D. dissertation, MIT.

Ladusaw, William A. 1988. Adverbs, negation and QR. In *Linguistics in the Morning Calm 2*, ed. by The Linguistic Society of Korea, 481-488. Seoul: Hanshin Publishing Co.

Lang, Ewald. 1979. Zum Status der Satzadverbiale. *Slovo a Slovesnost* 40: 200-213.

Law, Ann. 2003. Right dislocation in Cantonese as a focus-marking device. *UCL Working Papers in Linguistics* 15: 243-275.

Lebeaux, David. 1988. *Language acquisition and the form of grammar*. PhD dissertation, University of Massachusetts. Revised and extended edition, Amsterdam: John Benjamins, 2000.

Nuyts, Jan. 1993. Epistemic modal adverbs and adjectives and the layered representation of conceptual and linguistic structure. *Linguistics* 31:933-969.

Pesetsky, David and Esther Torrego 2007. The syntax of valuation and the interpretability of features. In *Phrasal and clausal architecture: Syntactic derivation and interpretation. In honor of Joseph E. Emonds*, ed. by Karimi, Simin, Vida Samiian and Wendy K. Wilkins, 262-294. Amsterdam: John Benjamins.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of Grammar*, ed. by Liliane Haegeman, 281-337. Dordrecht: Kluwer Academic Publishers.

Rooth, Mats. 1985. *Association with focus.* PhD dissertation, University of Massachusetts.

Rooth, Mats. 1996. Focus. In *The Handbook of Contemporary Semantic Theory*, ed. by Shalom Lappin. 271-297. Malden: Blackwell Publishers Ltd.

Shu, Chih-hsiang. 2006. The syntax of high adverbs: Overt and covert positions. Ms., Stony Brook University.

Svenonius, Peter. 2002. Subject position and the placement of adverbials. In *Subjects, Expletives, and the EPP*, ed. by Peter Svenonius, 201-242. New York: Oxford University Press.

Tancredi, Christopher D. 1990. Not only *even* but even *only*. Ms., MIT, Cambridge, Mass.

Watanabe, Akira. 2004. The Genesis of Negative Concord: Syntax and Morphology of Negative Doubling. *Linguistic Inquiry* 35.4: 559-612.

Children's Understanding of *dou* in Chinese

Peng Zhou
Macquarie University, Sydney

Abstract

The quantifier *dou* (roughly corresponding to English 'all') in Chinese has been the topic of much discussion in the theoretical literature. This study investigated children's knowledge of this quantifier, using a new methodological technique, which we dubbed Question-Statement Task. Two questions were discussed: (i) whether young Mandarin-speaking children know that *dou* is a universal quantifier that quantifies over the elements to its left, and (ii) whether they know that *dou* can quantify over *wh*-words. It was found that by age four Mandarin-speaking children already have the relevant knowledge. These results reflect early availability of adult-like linguistic knowledge of *dou*-quantification.

1. Introduction

The quantifier *dou* (roughly corresponding to English 'all') in Mandarin Chinese has been the topic of much discussion in the theoretical literature. Its syntax and semantics remain controversial. It is generally acknowledged, however, that *dou* is a universal quantifier that quantifies over the elements to its left. This study looked at Mandarin-speaking children's knowledge of this quantifier. To be specific, we investigated whether children know that *dou* is a universal quantifier.

The paper is organized as follows. First we introduce some basic facts about the quantifier *dou*. Then, we review previous research on Mandarin-speaking children's understanding of this quantifier. Finally, we present our study investigating Mandarin-speaking children's knowledge of this quantifier.

2. *Dou* as a Universal Quantifier

Despite the controversy on its syntax and semantics, there are some well-known characteristics of *dou*. In the remainder of this section, a basic overview of these characteristics is presented.

Dou is a universal quantifier, which quantifies over the elements to its left (see, e.g., Cheng, 1995; Lee, 1986; Pan, 2006). Consider the following two sentences, for example. In (1), *dou* quantifies over the individuals denoted by *tamen* 'they', thereby giving the sentence a universal reading. Thus sentence (1)

Peng Zhou

means 'each of them bought their own car.' However, with *dou* removed, as in (2), the sentence means 'the entire group of people denoted by *tamen* 'they' collectively bought a car'.

(1) Tamen dou   mai-le   yi-liang che.
    they   all   buy-ASP one-CL   car
    'They all bought a car.'

(2) Tamen mai-le   yi-liang che.
    they   buy-ASP one-CL car
    'They bought a car.'

Wh-words in Mandarin Chinese, such as *shei* 'who' and *shenme* 'what', can also be quantified over by *dou*. When this happens, they are no longer interpreted as interrogative words. Instead, they function roughly like variables that combine with *dou* to form a universally quantified NP (see Cheng, 1991, 1994; Huang, 1982; Li, 1992; Lin, 1996, 1998; for a detailed discussion of the interpretation of *wh*-words in Chinese).

The following sentences are used to illustrate.

(3) Shei dou xihuan ta.
    who all   like   he
    'Everyone likes him.'

(4) Yuehan shenme dou chi.
    John   what   all eat
    'John eats everything.'

In both (3) and (4), *dou* binds a *wh*-word to its left, *shei* 'who' in (3) and *shenme* 'what' in (4), thereby yielding a universally quantified NP, i.e., *everyone* as in (3) and *everything* as in (4). In this study, we are interested in how Mandarin-speaking children understand this quantifier *dou*, i.e., whether children know that *dou*, as a universal quantifier, can bind the *wh*-word in sentences like (3). We turn now to children's knowledge of this quantifier.

3. *Dou* in Child Mandarin

Few studies have been done to look at children's knowledge of *dou* in Mandarin Chinese. Only two studies, as far as we know, directly investigated children's knowledge of this quantifier. One is Lee's (1986) and the other is Hsieh's

(2008)[1]. Lee (1986) tested children's knowledge of *dou* in sentences like (5), in which *dou* quantifies over the subject noun phrase *xiongmao* 'panda' and thus the sentence means 'the pandas have all fallen asleep'. He tested children of six different age groups (i.e., 3-, 4-, 5-, 6-, 7-, and 8-year-olds), using a picture identification task, in which children were asked to choose from a pair of pictures the one that best corresponded to the test sentence.

(5) Xiongmao dou shuijiao-le.
    panda   all   sleep-ASP
    'The pandas have all fallen asleep.'

(6) Xiongmao shuijiao-le.
     panda   sleep-ASP
    'The pandas have fallen asleep.'

On a typical trial, children were presented with two pictures, one in which three pandas were sleeping, and the other in which two pandas were sleeping and a third panda was awake. Children were then presented with the test sentence (5). The findings were that only slightly more than half of the 3-year-olds correctly chose the picture in which three pandas were sleeping, but over 90% of the 4-year-olds chose the correct picture with three pandas sleeping. Based on the findings, Lee concluded that Mandarin-speaking children understand the universal quantificational feature of *dou* by age four.

There is a problem with this interpretation of the findings, however. The experiment lacked control trials in which the corresponding sentences without *dou* were tested against the same pairs of pictures. Consider sentence (6) without *dou*. Suppose that Mandarin-speaking children preferred the picture with the three pandas sleeping in response to sentence (6), just as they did for sentence (5). If so, this would undermine Lee's conclusion that children understand the meaning of *dou*. In the absence of such controls, we cannot exclude the possibility that children interpret sentences with *dou* as having the same meaning as their counterparts without *dou*. To determine whether or not children understand the universal quantificational feature of *dou*, both sentences with and without *dou* should be presented in the same contexts.

---

[1] There are two other studies done by Jia and her colleagues (1996, 1998), investigating how Mandarin-speaking children understand universal quantification, using sentences like *Suoyou de ren dou zai ban yi-ge xiangzi* 'All the men are carrying a box' and *Meiyi-ge ren dou zai ban yi-ge xiangzi* 'Every man is carrying a box'. Since their studies are not directly relevant to ours, the data of their studies will not be reported here. Readers who are interested are referred to Jia, Brooks and Braine (1996), and Brooks, Jia, Braine and Dias (1998).

Peng Zhou

Hsieh (2008) reported the data of a longitudinal study investigating Mandarin-speaking children's knowledge of scope. She collected the spontaneous speech of a Mandarin-speaking child from 1 year old up to late 4 years old and found that the child used *dou* as a universal quantifier when he was 4 years old. The following sentences from the child are used to illustrate.

(7) Shenme difang dou meiyou ba.
     what   place all   not   PART
     'Nowhere.'

(8) Shenme difang dou meiyou zhaodao.
     what   place all   not   find
     'I couldn't find it anywhere.'

In both (7) and (8), *dou* is used to quantify over the *wh*-phrase *shenme difang* 'what place', thereby rending it as the universal quantifier 'everywhere'. Thus with negation both statements mean 'nowhere'. The child used (7) to express 'he didn't go anywhere', and (8) to express 'he cannot find the thing at any place'. Based on the data, it seems that children use *dou* as a universal quantifier by age four. However, this study only recorded the data of one child. We cannot be sure to what extent the data of this child can be generalized to all the Mandarin-speaking children of the same age. More subjects need to be tested before we can make a generalization.

To summarize, despite the problems they may have, previous studies seem to suggest that children know that *dou* is a universal quantifier by age four. In this study, we want to confirm whether Mandarin-speaking children analyse *dou* as a universal quantifier by age four, using sentences like (3).

4. Our Study

In our experiment, we investigated whether children know that *dou* is a universal quantifier. The experiment contrasted minimal pairs of sentences with *dou* and ones without *dou*, as illustrated in (9) and (10), respectively.

(9) Shei dou meiyou pa-shang dashu.
    who all   not   climb-up big tree
    'Everyone didn't climb up the big tree.'

(10) Shei meiyou pa-shang dashu?
     who not   climb-up big tree
     'Who didn't climb up the big tree?'

In (9), *dou* quantifies over the *wh*-word *shei* 'who', thereby rendering it as a universal quantifier 'everyone'. Thus (9) is a negative statement, meaning 'nobody climbed up the big tree'. Without *dou*, the *wh*-word *shei* 'who' in (10) marks the sentence as a question, asking 'who didn't climb up the big tree?' So the hypothesis is that if children know that *dou* is a universal quantifier that quantifies over the elements to its left, then they should interpret sentences like (9) as statements and ones like (10) as questions. But if, on the other hand, children do not have the relevant knowledge of *dou*, then they might be expected to interpret sentence (9) as a question, just like sentence (10).

4.1 Subjects

We tested 30 Mandarin-speaking children between the ages of 3;5 and 4;9 (mean age 4;2). They were recruited from the kindergarten at Beijing Language and Culture University. In addition, 30 Mandarin-speaking adults were tested as controls, all postgraduate students at Beijing Language and Culture University.

4.2 Method and Procedures

In order to evaluate the experimental hypothesis, we designed a new methodological technique, which we dubbed Question-Statement Task. The task involves two experimenters. One acted out stories using toy characters and props, and the other played the role of a puppet who watched the stories alongside the subject. After each story, the puppet attempted to tell the subject what he thought had happened in the story, using a test sentence. However, sometimes the puppet didn't pay close attention to the story and thus was not sure about what happened in the story. If this was the case, then the puppet would ask the child a question, using a test sentence. On each trial, the subject's task was to decide whether the puppet accurately said what happened in the story or asked a question about the story. Whenever the puppet said what happened in the story (if the subject thought so), the subject was instructed to judge whether the puppet was right or wrong. But if the puppet asked a question about the story (if the subject thought so), the subject was instructed to answer the question and then ask the puppet to pay closer attention to the next story.

The subjects were introduced to the task individually and then tested individually. In order to familiarize the subjects with the task, they were given two practice trials before the actual test, one in which the puppet informed the subject what happened in the story and one in which the puppet asked the subject a question. Only those subjects who correctly judged whether the puppet made a

statement about the story or asked a question about the story were included in the actual test.

4.3 Materials

There were 6 test stories and for each story, two types of test sentences were created: one with *wh*-word + *dou* as in (9), and one with *wh*-word alone as in (10). An example is given as follows.

Three dogs (a black dog, a white dog and a brown dog) are going to have a tree climbing contest. They are all very good at tree climbing. This time, they need to climb a big tree and a small tree. They start with the small tree. They all made it to the top easily, as illustrated in Figure 1. Then they come to the big tree. It is much taller than the small tree. The black dog is really a good climber. He touches the top of the tree easily. But the white dog and the brown dog have troubles getting into the branches. Each time they lift their front paws, their back paws slide off the branches. No luck, they didn't climb to the top. They failed. Figure 2 illustrates the last scene at the story.

Fig.1. The first half of the story, Experiment 1

Fig.2. The second half of the story, Experiment 1

Since the test sentences involve negation, it is important that our test stories satisfy the felicity conditions associated with the use of negation. Crain et al. (1996) proposed the Condition of Plausible Dissent. This condition is based on Russell's (1948) observation that a negative judgement is appropriate only when the correlative positive judgement has already been made or considered. In our task, subjects were asked to say whether the test sentences were true or false if they thought the puppet told them what happened in the story. Following Russell's observation, it is appropriate to ask the subjects for a negative judgement of a sentence only if the corresponding positive judgement has been under consideration at some point of the story. In order to satisfy this condition, the puppet produced a positive lead-in before the test sentence, which corresponded to the first half of the story. In this story, the positive lead-in was *San-zhi gou dou pa-shang-le xiaoshu* 'All the three dogs climbed up the small tree.' After the positive lead-in, the puppet either produced the test sentence, as in (9) or the test sentence, as in (10), repeated here as (11) and (12), respectively.

(11) Shei dou meiyou pa-shang dashu.
     who all   not   climb-up big tree
     'Everyone didn't climb up the big tree.'

(12) Shei meiyou pa-shang dashu?
     who not   climb-up big tree
     'Who didn't climb up the big tree?'

Peng Zhou

The subjects were then divided into two groups. One group (15 chid subjects and 15 adult subjects) heard the puppet produce the test sentence, as in (11). And the other group (15 child subjects and 15 adult subjects) heard the puppet produce the test sentence, as in (12). There were altogether six test stories. For each story, one group of subjects was presented with one type of the test sentences and the other group was presented with the other type of the test sentences. And the two types of the test sentences were counterbalanced across the two groups. That is, each group of subjects was presented with both types of test sentences and there were an equal number of the two types of sentences for both groups of subjects.

In addition to the test trials, each subject witnessed four control trials. On two of the control trials, the puppet produced the questions in (13) and (14). These two trials were used to see whether subjects understand simple *wh*-questions. On the other two trials, the puppet produced simple statements with *dou*, as in (15) and (16). These two trials were included to verify that subjects could understand simple statements with *dou*. And they were also used to counterbalance the "yes" and "no" answers throughout the trials. The statements in the test trials were all false in the relevant stories, so we made the statements in (15) and (16) true in the relevant stories. The test and control trials were presented in a pseudo-random order.

(13) Shei zhuangdao-le huluobo?
     who   hit-ASP   carrot
     'Who hit the carrot?'

(14) Shei nadao-le   beike?
     who get-ASP   shell
     'Who got the shell?'

(15) Xiaomao he xiaotuzi dou chi-le   yu.
     cat   and bunny   all eat-ASP fish
     'The cat and the bunny both ate a fish.'

(16) Xiaonanhai he xiaonühai dou qi-le   changjinglu.
     boy   and girl   all ride-ASP giraffe
     'The boy and the girl both rode a giraffe.'

Before we report the results, let's turn to our hypothesis just discussed. We propose that if children know that *dou* is a universal quantifier that quantifies over the elements to its left, then they should interpret sentences with *wh*-word + *dou* as negative statements and ones with *wh*-word alone as questions. In the given story, they should reject the puppet's statement (11), by pointing out that

the black dog climbed up the big tree, and they should provide an answer to the puppet's question (12), by pointing out that the white dog and the brown dog didn't climb up the big tree. But, if on the other hand, children do not have the relevant knowledge of *dou*, then they might be expected to interpret sentences with *wh*-word + *dou* as having the same meaning as their counterparts with *wh*-word alone. In the given story, children should be expected to respond similarly to both (11) and (12), i.e., pointing out that the white dog and the brown dog didn't climb up the big tree.

4.4 Results and Discussion

We recorded the responses of the subjects to the two types of sentences produced by the puppet. All the subjects responded correctly to the control trials.

It was found that children responded to test sentences with *wh*-word + *dou* by rejecting the puppet's statements 96.3% of the time[2] and adults rejected them 100% of the time. There was no significant difference between children and adults ($Z = 1.86$, $p = .1$). When asked to justify their rejections, both children and adults pointed out that one of the characters did perform the action mentioned in the test sentences. In the given story, they rejected the puppet's statement (11) by making reference to the fact that the black dog climbed up the big tree. By contrast, both children and adults provided an answer in response to test sentences with *wh*-word alone, pointing out that the other two characters didn't perform the action mentioned in the test sentences (children: 95.1%[3] vs. adults: 100%; $Z = 1.86$, $p = .1$). In the given story, they responded to the puppet's question (12) by pointing out that the white dog and the brown dog didn't climb up the big tree.

The results of this experiment clearly show that children know that *dou* turns sentences like (11) into statements, in contrast to their counterparts without *dou* like (12), which children correctly interpreted as questions. This is compelling evidence that children know that *dou* is a universal quantifier by age four. These results confirm the conclusion reached by previous research that Mandarin-speaking children analyse *dou* as a universal quantifier by age four.

---

[2] For the other 3.7% of the time, children said "yes" to the puppet's statements. But none of the child subjects responded to this type of test sentences by providing an answer as they did for test sentences with *wh*-word alone.

[3] For the other 4.9% of the time, children gave a wrong answer to the test sentences, which we think might be due to their distraction during the story. But none of the child subjects responded to this type of test sentences in the same way as they did for test sentences with *wh*-word + *dou*.

Peng Zhou

References

Brooks, P.J., Jia, X.-D., Braine, M.D.S., and Dias, M.D.G. 1998. What do children know about the universal quantifiers *all* and *each*? *Cognition* 60: 235-268.

Cheng, L. L.S. 1991. On the typology of wh-questions. Ph.D. Dissertation, MIT.

Cheng, L. L.-S. 1994. Wh-words as polarity items. *Chinese Languages and Linguistics* 2: 615-640.

Cheng, L. L.-S. 1995. On *dou*-quantification. *Journal of East Asian Linguistics* 4: 197-234.

Crain, S., Thornton, R., Boster, C., Conway, L., Lillo-Martin, D., and Woodams, E. 1996. Quantification without qualification. *Language Acquisition* 5: 83-153.

Hsieh, M.-L. 2008. Acquiring scope: A longitudinal study. *Taiwan Journal of Linguistics* 6 (1): 55-96.

Huang, C.-T. J. 1982. Logical relations in Chinese and the theory of grammar. Ph.D. Dissertation, MIT.

Jia, X.-D., Brooks, P.J., and Braine, M.D.S. 1996. A study of Chinese children's comprehension of universal quantifiers. In Clark, E. (ed.), *Proceedings of the 27th Annual Child Language Research Forum*, 167-173. CA: CSLI, Stanford.

Lee, T. H.-T. 1986. Studies on quantification in Chinese. Ph.D. dissertation, University of California, Los Angeles.

Li, Y.-H.A. 1992. Indefinite wh in Mandarin Chinese. *Journal of East Asian Linguistics* 1: 125-155.

Lin, J.-W. 1996. Polarity licensing and wh-phrase quantification in Chinese. Ph.D. Dissertation, University of Massachusetts, Amherst.

Lin, J.-W. 1998. On existential polarity wh-phrases in Chinese. *Journal of East Asian Linguistics* 7: 219-255.

Pan, H.-H. 2006. Focus, tripartite structure, and the semantic interpretation of Mandarin *dou*. *Research and Exploration on Grammar* 13: 163-184

Russell, B. 1948. *Human knowledge: Its scope and limits*. London: Allen and Unwin.

Children's Interpretation of *Before* and *Or* in English and Mandarin Chinese

Anna Notley, Peng Zhou, Britta Jensen, Stephen Crain
Macquarie University, Sydney

Abstract

This study investigates children's interpretation of disjunction (*or / huozhe*) in sentences with English *before* and Mandarin *zai…zhiqian*. In English, when disjunction appears in the scope of *before* it gives rise to a conjunctive interpretation (*e.g., The dog reached the finish line before the turtle or the bunny* means 'The dog reached the finish line before the turtle <u>and</u> before the bunny'). This meaning is a logical consequence of the semantics of the word *or* being acted upon by the semantics of the word *before*. However, in Mandarin Chinese, disjunction can take scope over *zai…zhiqian*. When this happens, disjunction is no longer acted on by the temporal conjunction, and no conjunctive interpretation arises (*e.g., The dog reached the finish line before the turtle or the bunny* can mean 'The dog reached the finish line before the turtle <u>or</u> it reached the finish line before the bunny (but I don't know which one)'). If children are guided by the adult input in their acquisition of *before...or* sentence meanings, then they should show different interpretations across languages: English children should understand the example sentence to mean the dog reached the finish line first, while Mandarin children could understand the sentence to mean the dog reached the finish line second. On the other hand, if children are guided by innate logical principles during the acquisition process, both groups of children should compute the conjunctive interpretation. Our study was designed to adjudicate between these alternative learning scenarios.

1. Introduction

This study investigates children's interpretation of disjunction (*or / huozhe*) in sentences with English *before* and Mandarin *zai…zhiqian* (*e.g., The dog reached the finish line before the turtle or the bunny*). The behaviour of disjunction in sentences with this temporal conjunction differs in these two languages, providing an interesting testing ground for theories of how children acquire compositional semantic rules.

To frame our study, we will begin with a discussion of the interpretation of disjunction in sentences containing logical words like *not, none* and *every*. We

Anna Notley, Peng Zhou, Britta Jensen, Stephen Crain

then introduce some cross-linguistic differences in how such sentences are interpreted, and we review previous research on children's interpretation of disjunction in various languages. The results of previous studies indicate that children are guided by logical principles in their interpretation of disjunction in sentences with negative operators like *not* and *none*. However, to our knowledge no work has been done on the non-negative operator *before*. The present study asks whether children across languages also adhere to logical principles in interpreting disjunction in sentences with *before*.

1.1 Downward Entailment and the Conjunctive Interpretation of Disjunction

There is a class of words called *downward entailing* (DE) operators which encompasses a wide range of parts of speech in natural language. For example, the negative operator *not*, determiners like *every* and *none*, the preposition *without* and the conjunction *before* are all downward entailing operators. They can be treated as a natural class because they share several unifying properties. Firstly, they license inferences from general terms (*e.g.,* Romance languages) to specific terms (*e.g.,* French). For example, upon hearing the statement *John did not learn a Romance language*, we can validly infer that 'John did not learn French' (or any other specific Romance language); similarly, if the statement *Dinosaurs lived before modern mammals* is true, then it must be true that 'Dinosaurs lived before foxes', and if *Every student took the bus to university* is true, then 'Every first year student took the bus to university'. The licensing of inferences from sets to their subsets is the defining property of DE operators.

A second property of downward entailing operators is that they license a conjunctive interpretation of disjunction. We will use the example of negation to illustrate. Consider the English sentence *I do not like broccoli or cauliflower*. This statement is understood as meaning 'I do not like broccoli <u>and</u> I do not like cauliflower'. This is called the conjunctive interpretation of disjunction, and is, in fact, the logical consequence of the disjunction operator *or* in English having an inclusive meaning. Inclusive *or* is associated with three truth conditions. The statement 'P *or* Q' is true if either P is true (but Q is not), or Q is true (but P is not), or both P and Q are true. In other words, the statement 'P *or* Q' is false in just one situation: when neither P nor Q is true. In the scope of negation, the truth conditions for *or* are reversed. So, '*not* (P *or* Q)' is true precisely when 'P *or* Q' is false, that is when neither P nor Q is true. This relationship is captured in one of de Morgan's laws of propositional logic (where '$\neg$' symbolizes 'not', '$\vee$' symbolizes 'or', and '$\wedge$' symbolizes 'and'):

(1) ¬(P ∨ Q) ⇒ ¬P ∧ ¬Q   (not P or Q is logically equivalent to not P and not Q)

Interestingly, in natural languages, the conjunctive interpretation of disjunction not only arises when disjunction is negated, but when it appears in downward entailing environments in general. For example, when *or* appears in sentences with *none, without, at most, every* or *before*, it generates a conjunctive interpretation, as illustrated in (2).

(2) (a) *None of the students took maths or biology*
⇒ None of the students took maths and none of the students took biology

(b) *I left the restaurant without my purse or my camera*
⇒ I left the restaurant without my purse and I left the restaurant without my camera

(c) *At most three researchers travelled by plane or by train to the conference*
⇒ At most three researchers travelled by plane and at most three researchers travelled by train to the conference

(d) *Every passenger who ate the chicken or the beef on the plane was ill*
⇒ Every passenger who ate the chicken on the plane was ill and every passenger who ate the beef on the plane was ill

(e) *Jane arrived at the pool before Mary or Sue*
⇒ Jane arrived at the pool before Mary and Jane arrived at the pool before Sue

Some of these operators clearly contain negation as part of their meaning (*e.g., none, without* and *at most*, which is equivalent to *no more than*). In view of the logical relationship between negation and disjunction, it is clear why these operators trigger a conjunctive interpretation of disjunction. However, some downward entailing operators don't fit the pattern established by de Morgan's laws, in that they do not require a negative meaning component to license the conjunctive interpretation of disjunction. *Every* and *before* are two such operators.

In the case of these operators, it appears that a different logical process gives rise to the conjunctive interpretation of disjunction. This process is based on set relations, as they apply to the universal quantifier. For example, in (2d) *or* is used to define subsets of a universally quantified superset ('passengers who ate chicken' and 'passengers who ate beef' are subsets of the overall set of

Anna Notley, Peng Zhou, Britta Jensen, Stephen Crain

'passengers'). So, when we universally quantify over 'passengers', the resulting superset necessarily contains both passengers who ate chicken <u>and</u> passengers who ate beef (as well as any passengers who ate both chicken and beef). In other words, the conjunctive interpretation of disjunction arises in this case because, when assessing the overall truth of the universally quantified statement, all 3 truth conditions for inclusive *or* must be true. Recall that in the case of negatively flavoured downward entailing operators, by contrast, we are considering the one truth condition for *or* that is normally false.

The conjunctive interpretation of disjunction in sentences with *before* can also be accounted for by a logical subset-superset relationship introduced by the universal quantifier. This follows from the analysis of the semantics of this operator as quantifying over time points. That is, when an event A occurs *before* an event B, then at least some time point in event A occurred before *every* time point in event B (Anscombe, 1964, Heinamaki, 1972). This hidden universal quantification within the conjunction *before* is presumably the source of the conjunctive interpretation of disjunction. For example, in (2e), event A is 'Jane's arrival at the pool' and event B is 'Mary's arrival at the pool <u>or</u> Sue's arrival at the pool'. The disjunction has effectively split event B into two 'subsets'. If Jane's arrival at the pool (event A) has occurred before event B, then it must have occurred before every time point in event B – every time point in the event of Mary's arrival, <u>and</u> every time point in the event of Sue's arrival (or alternatively every time point in the event of the simultaneous arrival of both girls).[1]

1.2 Cross-Linguistic Differences in Downward Entailment Properties

For the conjunctive interpretation of disjunction to arise in a sentence containing a downward entailing (DE) operator and disjunction, the disjunction operator must be interpreted <u>within</u> the semantic scope of the DE operator. We will now review some interesting cross-linguistic differences in the syntactic environments which give rise to these necessary scope relations, again starting with the example of negation.

In some languages, like English, disjunction is interpreted within the scope of negation in both simple and complex clauses, as in (3) and (4) below.

---

[1] Note that in this case, there is little difference between the effect of *or* versus the conjunction operator *and* (e.g. *I arrived at the pool before Mary and Sue)*. Conjunction also effectively divides event B into two subset events, which then must each be considered in assessing the overall truth of the statement. This is in contrast to universally quantified statements like '*Every passenger who ate the chicken and the beef on the plane was ill'* which means we only have to consider whether the overlapping portion of the set of chicken-eaters and the set of beef-eaters fell ill, in order to assess the truth of the statement.

(3)   *John does not like broccoli or cauliflower*
       ⇒ 'John does not like broccoli and John does not like cauliflower'

(4)   *I do not think John likes broccoli or cauliflower*
       ⇒ 'I do not think John likes broccoli and I do not think John likes cauliflower'

Languages which behave like English in this respect include German, French, Greek, Romanian, Bulgarian, and Korean (Szabolcsi, 2002).

In some other languages, however, such as Mandarin, the conjunctive interpretation of disjunction only arises in <u>complex</u> clauses, where negation appears in a higher clause than the clause that contains disjunction, as illustrated in (5).

(5)   *Wo   bu   renwei ta   xihuan xilanhua   huozhe huayecai*
        I   not   think   he   like   broccoli   or   cauliflower
        'I do not think he likes broccoli or cauliflower'
        ⇒ 'I do not think he likes broccoli and I do not think he likes cauliflower'

When both negation and disjunction appear in the same clause, as in (6), the disjunction operator tends to be interpreted as taking scope over negation. So, in Mandarin, a sentence like 'John does not like broccoli or cauliflower' can easily mean 'It is broccoli or cauliflower that John doesn't like'.[2]

(6)   *Ta   bu   xihuan xilanhua   huozhe huayecai*
        He   not   like   broccoli   or   cauliflower
        'He does not like broccoli or cauliflower'
        ⇒ 'It is broccoli or cauliflower that he doesn't like'

Languages which allow disjunction to be interpreted as taking scope over local negation in this way include Hungarian, Japanese, Russian, Serbo-Croatian, Slovak, and Polish (Goro and Akiba 2004a, Goro and Akiba 2004b, Szabolcsi 2002). Note that, as a result of this scoping behaviour, the reading of disjunction in such sentences in these languages is typically exclusive (*e.g.,* 'It is either broccoli or cauliflower (but not both) that he doesn't like'). This is because hearers may compute a scalar implicature for the sentence meaning at a pragmatic level. This scalar implicature is based on a scale formed by the operators *or* and *and*. The ordering of operators on the scale is determined by information strength, where a term α is 'stronger' than another term β if α

---

[2] Note that the notion of scope under consideration here is not related to the linear ordering of the words in the sentence. Negation occurs before disjunction in Mandarin, however disjunction can be interpreted taking scope over negation.

Anna Notley, Peng Zhou, Britta Jensen, Stephen Crain

asymmetrically entails β. Since the truth conditions of *and* are a subset of the truth conditions of inclusive *or*, statements with *and* asymmetrically entail the corresponding statements with *or,* so *and* is the stronger term on the scale. Following the Gricean conversational maxim of quantity 'to make your contribution as informative as is required and not more informative than is required' (Grice, 1975), a hearer will generally assume that if a speaker uses *or*, they are not in a position to use the stronger term *and* to describe the situation under consideration. Hearers therefore remove the truth conditions associated with *and* from the meaning of *or*, giving the disjunctive statement an exclusive reading (Horn, 1996).

We have seen that the behaviour of disjunction in sentences with local negation differs across languages. We can now ask whether languages differ in the behaviour of disjunction in sentences with the universal quantifier. In answering this question it is important to point out that the conjunctive interpretation of disjunction only arises in what we call the restrictor of the universal quantifier, the noun phrase to which it is bound syntactically (*e.g., Every* [*passenger who ate chicken or beef*]$_{Restrictor}$ *on the plane was ill*). When disjunction occurs in this position, it becomes part of the constituent headed by *every*.[3] Accordingly, as one would expect given this close syntactic relationship, there does not seem to be a cross-linguistic difference in how sentences of type (2d) are interpreted in English-type and Mandarin-type languages (at least in the Mandarin-type languages we have reviewed: Mandarin, Japanese, and Hungarian); the conjunctive interpretation of disjunction arises in both language types in the restrictor of *every*. But how does disjunction behave across languages in sentences with *before*? Here, the conjunctive interpretation of disjunction presumably arises because of the action of the universal quantifier within the semantics of *before*, yet at the same time disjunction is not formally in the syntactic restrictor of *every*. Interestingly, in this context, we <u>do</u> find a difference in how disjunction is interpreted in English as compared to some other languages like Japanese, Hungarian, Danish and Mandarin. For example, in Mandarin, disjunction can scope over *zai...zhiqian,* as illustrated in (7).

(7) *Jian   zai Mali huozhe Su   zhiqian   dao-le   shuichibian*
     Jane   at Mary   or   Sue before   arrive-ASP pool-side

---

[3] By contrast, when disjunction occurs outside the restrictor (e.g. *Every* [*plane passenger who was ill*]$_{Restrictor}$ *ate chicken or beef*), any of the range of truth conditions of disjunction will make the sentence true. That is, the sentence *Every plane passenger who was ill ate chicken or beef* is true if every passenger who was ill ate chicken, <u>or</u> if every passenger who was ill ate beef, <u>or</u> if some of the ill passengers ate chicken and some ate beef. Only one of these scenarios need be true for the whole sentence to be true.

'Jane arrived at the pool before Mary or Sue'
$\Rightarrow$ 'It is before Mary or Sue that Jane arrived at the pool'

Cross-linguistic scoping differences like these provide a prime testing ground for theories of language acquisition. Firstly, we can ask how children learning an English-type language interpret disjunction in sentences with local negation (or other negatively flavoured DE operators), compared to children learning a Mandarin-type language. An input-based model of language acquisition predicts that, within a given language, children's responses should mirror the adult input. So, we would expect to see children learning an English-type language assign the conjunctive interpretation to disjunction in the scope of local negation, but children learning a Mandarin-type language to allow disjunction to scope over local negation. By contrast, some researchers have suggested that children draw on universally innate concepts about the meaning of disjunction and its logical interaction with the class of DE operators (Crain *et al.* 2005, Crain *et al.* 2006, Crain and Thornton 2006). According to such an innateness hypothesis, we might expect to see children across languages initially compute the conjunctive interpretation of disjunction in the scope of negation, regardless of adult input.

Secondly, we can ask whether children's patterns of interpretation of disjunction extend beyond just negatively flavoured DE contexts to encompass universally quantified DE contexts as well. This study addresses this question by testing children's interpretation of disjunction in sentences with *before*. The findings of the study are of particular interest because an input-based model of language acquisition would seemingly have considerable difficulty making any predictions at all about children's interpretations of such sentences, as the relevant sentences occur very rarely in the input. For example, in a survey of 224,797 parental utterances in 7 English corpuses on CHILDES,[4] we found only 2 instances of a construction containing disjunction combined with *before* (*e.g.,* from the MacWhinney corpus: *I'll wait till Mark comes in before I read anything or do anything okay?*). We turn now to a brief review of previous research in this area, before outlining our own methodology.

1.3 Previous Child Research on Downward Entailment Relations

In 2002, Crain, Gardner, Gualmini, & Rabbin showed that 3- to 5-year old English-speaking children, like adults, consistently assign a conjunctive interpretation to disjunction when it appears in the scope of local negation. They presented 30 children with two different test sentence types. In both types of

---

[4] The MacWhinney corpus, the Brown corpus (Adam, Eve and Sarah), and the New England corpus (Folders 14, 20, and 32). These corpuses record input to children from age 1-7.

Anna Notley, Peng Zhou, Britta Jensen, Stephen Crain

sentences, negation preceded disjunction. However, in one type negation locally c-commanded the disjunction operator (*e.g., The girl who stayed up late will not get a dime or a jewel*). This results in a conjunctive interpretation for adult speakers (*e.g.,* 'The girl who stayed up late will not get a dime <u>and</u> the girl who stayed up late will not get a jewel'). In the other type of test sentence, negation did not c-command the disjunction operator (*e.g .The girl who did not go to sleep will get a dime or a jewel*). This does not result in a conjunctive interpretation (*e.g.,* 'The girl who did not go to sleep will get a dime <u>or</u> the girl who did not go to sleep will get a jewel'). Children were tested using a truth value judgment task in which a story about two girls waiting for the tooth fairy was acted out in front of them and a puppet. At the end of the story, it turned out that the girl who stayed up late (i.e. the girl who did not go to sleep) got a jewel. This context made those sentences false in which negation c-commanded disjunction, whereas it made those sentences true in which negation did not c-command disjunction. The child subjects were sensitive to this, judging sentences like *The girl who stayed up late will not get a dime or a jewel* to be false 92% of the time, and judging sentences like *The girl who did not go to sleep will get a dime or a jewel* to be true 87% of the time. This result was replicated by Gualmini & Crain in 2005 (Gualmini 2005, Gualmini and Crain 2005), and has also been shown to hold in child English for the operator *none* (Gualmini and Crain 2002).

What about children learning a language in which disjunction can be interpreted as taking scope over negation in simple negative statements? In 2004, Goro & Akiba tested 30 3- to 6- year-old Japanese-speaking children on sentences like *The pig did not eat the carrot or the pepper* in contexts in which it turned out that the pig in question ate a pepper. Whereas English speakers judge such sentences to be false, Japanese adults judged the corresponding Japanese sentences to be true. This is because the interpretation of the sentence by Japanese-speakers allows disjunction to take scope over negation. So the sentence corresponding to *The pig did not eat the carrot or the pepper* can be paraphrased as 'It is either a carrot or a pepper that the pig did not eat'. Since the pig did not eat a carrot, Japanese-speaking adults judged the sentence to be true. However, the Japanese-speaking children that were tested by Goro and Akiba differed markedly from adults. Children judged such sentences to be false 75% of the time. Four of the oldest children were effectively adults and consistently accepted the test sentences. When the results of these four children were removed, the rejection rate for the remaining 26 children was 87%. It appears then that Japanese-speaking children initially compute the conjunctive interpretation for disjunction in the scope of local negation, unlike Japanese-speaking adults. These findings are difficult to explain on an input-based model

of language acquisition. Instead, they favour the hypothesis that children across languages draw upon possibly innate universal logical concepts about the meaning of disjunction and its interaction with downward entailing operators.

To further test this hypothesis, other studies have sought to find out how children learning different languages respond to disjunction in the scope of downward entailing operators other than negation. Some work has been done in English and Mandarin on children's interpretation of disjunction in the restrictor of the universal quantifier, which, as we have discussed, gives rise to a conjunctive interpretation in both languages. For example, a sentence like *Every troll who ordered French fries or onion rings got mustard* entails that 'Every troll who ordered French fries got mustard <u>and</u> every troll who ordered onion rings got mustard'. It has been shown that 3- to 5-year-old English-speaking and Mandarin-speaking children consistently reject sentences of this type in contexts in which, for example, only trolls who ordered French fries got mustard. Moreover, children learning both languages distinguish between the downward entailing restrictor of *every* and the non-downward entailing nuclear scope of *every* in sentences like *Every ghostbuster will choose a cat or a pig*. Both English-speaking and Mandarin-speaking children consistently accept sentences like this in contexts in which, for example, ghostbusters choose cats or pigs, but not both (Boster and Crain 1993, Gualmini *et. al.* 2003, Su and Crain, forthcoming). These findings are in line with the hypothesis that children across languages draw upon innate universal logical concepts about the meaning of disjunction and its interaction with downward entailing operators.

Even stronger support for the innateness hypothesis could come from investigations of children's interpretation of disjunction in the scope of a non-negative downward entailing operator in which cross-linguistic differences arise. One such operator, as illustrated above, is the temporal conjunction *before*. The present study thus investigates how English-speaking and Mandarin-speaking children interpret *or* and *huozhe* in the scope of *before* and *zai…zhiqian* respectively.

2. Predictions

Recall that in an English sentence, the downward entailing operator *before* licenses a conjunctive interpretation of disjunction, as in (2e), repeated here as (8).

(8)   *Jane arrived at the pool before Mary or Sue*
          $\Rightarrow$ Jane arrived at the pool before Mary and Jane arrived at the pool before Sue

Anna Notley, Peng Zhou, Britta Jensen, Stephen Crain

By contrast, in Mandarin, disjunction can take scope over a local downward entailing operator like *zai…zhiqian* as in (7), repeated here as (9).

(9) *Jian   zai Mali huozhe Su   zhiqian dao-le   shuichibian*
     Jane   at Mary or   Sue before arrive-ASP   pool-side
     'Jane arrived at the pool before Mary or Sue'
     ⇒ 'It is before Mary or Sue that Jane arrived at the pool'

The difference between the two languages is, however, not as clear-cut as in the case of negation and disjunction. That is, a Mandarin speaker may also compute the conjunctive interpretation of disjunction in the scope of *zai…zhiqian,* just as in English. Nonetheless, a reading in which disjunction is interpreted as taking scope over *zai…zhiqian* in sentences like (9) is much more accessible to Mandarin-speakers than the corresponding reading with disjunction and *before* for English speakers. When disjunction is interpreted as taking scope over *zai…zhiqian*, the reading that results for Mandarin-speakers can engage an implicature of exclusivity (*e.g.,* 'It is either before Mary or before Sue (but not before both) that Jane arrived at the pool'). Such a reading is, at best, a faint possibility in English, and requires a particularly marked prosodic contour in which there is a long pause before disjunction, with heavy stress placed on the word *or*.[5]

Given these differences between English and Mandarin, if children rely on the scope relations attested in their local language, then we would expect English-speaking children to compute the conjunctive interpretation of disjunction in sentences with *before*. On the other hand, we would expect Mandarin-speaking children to show a mix of both the conjunctive and non-conjunctive interpretation of disjunction in sentences with *zai…zhiqian*. By contrast, if children adhere to a logical set relation principle dictated by the hidden universal quantifier in the semantics of *before*, then they are expected to initially compute the conjunctive interpretation of disjunction in sentences with *before* in English, and in sentences with *zai…zhiqian* in Mandarin, regardless of the scope relations attested in the target language. The present study was designed to evaluate the different predictions of these alternative accounts of language development.

---

[5] Which raises the possibility that in such cases an ellipsis of a longer sentence is being performed (*e.g.* 'Jane arrived at the pool before Mary…or (she arrived at the pool) before Sue')

3. Study

3.1 Subjects

We tested 24 English-speaking children between the ages of 3;4 and 5;1 (13 boys, 11 girls, mean age 4;4) and 20 Mandarin-speaking children between the ages of 4;6 and 5;4 (mean age 4;7). The English-speaking children were recruited from two daycare centres at Macquarie University, Sydney, Australia and all had English as their sole home language. The Mandarin-speaking children were recruited from the kindergarten at Beijing Language and Culture University in China and all had Mandarin as their sole home language. In addition we tested 20 English-speaking undergraduate students at Macquarie University (aged 18-27, mean age 21), and 20 Mandarin-speaking postgraduate students at Beijing Language and Culture University (aged 25-30, mean age 27).

3.2 Methodology

The subjects were tested using a truth value judgement task. This research technique is designed to investigate which meanings children can and cannot assign to sentences (Crain and Thornton 1998). The task involves two experimenters – one acting out stories with toy characters and props, and the other playing the role of a puppet who watches the stories alongside the child. At the end of each story, the puppet explains to the child subject what he thinks happened in the story. The child's task is to decide whether the puppet said the right thing or not. If the child informs the puppet that he was wrong, then the child is asked to explain to the puppet what really happened.

Four racing stories were devised, each with three participants. In each race, one participant came first, one second, and one last. At the end of the race, the participants were placed on a three-tiered podium to reflect the order they had come in (first, second, or third), serving as a reminder to the child of the events of the story. The puppet then produced a test sentence, such as (10).

(10)   (a) The dog reached the finish line before the turtle or the bunny

        (b) *Xiaogou zai wugui huozhe tuzi zhiqian paodao-le zhongdian*
           dog   at   turtle   or   rabbit before reach-ASP finish line
           'The dog reached the finish line before the turtle or the bunny'

Two of the four test sentences described contexts in which the referent of the subject NP (*e.g.,* the dog) came first. We will call this the *First-Place condition*. The other two test sentences described contexts in which the referent of the subject NP came second. We will call this the *Second-Place condition*. We expected that if children computed a conjunctive interpretation of disjunction,

Anna Notley, Peng Zhou, Britta Jensen, Stephen Crain

they should judge (10) to be a true description of stories in the First-Place condition, but a false description of stories in the Second-Place condition. On the other hand, if children allowed disjunction to take scope over *before*, then they should judge (10) to be true in both conditions (i.e. when the dog comes first, before both other participants, as well as when the dog comes second, before only one other participant).

It was important to ensure that child subjects were actually processing both disjuncts when judging the test sentences. To verify this, we ordered the disjuncts so that if the children made a false judgment in the Second-Place condition we could be sure they were responding to the full test sentence. That is, the first disjunct always referred to the participant who had come last, while the second disjunct referred to the participant who had come first. For example, in our swimming race story, a horse, a duck and a dolphin each had to swim to a shell at the end of a pool. The dolphin came first in this race, the duck second, and the horse last. After the story, the puppet said: *[The duck]*$_{2nd\ place}$ *got his shell before [the horse]*$_{3rd\ place}$ *or [the dolphin]*$_{1st\ place}$. Children could only reject this statement if they processed both disjuncts, and they computed a conjunctive interpretation: it was not true that the duck got his shell before the horse <u>and</u> before the dolphin. On the flip side, children might agree with the puppet's statement for two reasons – either they allowed disjunction to take scope over *before* (it was true that the duck either came before the horse <u>or</u> before the dolphin), or they simply only processed the first disjunct (it was also true that the duck got his shell before the horse). To make sure that any 'true' judgments in the Second-Place condition stories were genuinely due to children allowing disjunction to take scope over *before,* we designed a control story. This control was identical to the test scenarios in that three participants took part in a race, but the control sentence contained *and* instead of *or* as given in (11). In the story, Tiger came first, followed by the pig, and then the elephant. To successfully reject this control sentence, children had to be processing both disjuncts.

(11)   (a) The pig jumped to the finish line before the elephant and tiger
       (b) *Xiaozhu   zai   daxiang   he   tiaotiaohu   zhiqian*
              pig   at   elephant   and   Tigger   before
            tiaodao-le   zhongdian
             jump-to-ASP   finish line
            'The pig jumped to the finish line before the elephant and Tiger.'

Each test sentence was followed by a filler sentence which contained neither *before* nor disjunction (*e.g., In that race, the turtle fell over*). The fillers allowed us to balance the total number of true and false statements, and check whether the

children had been paying attention to the stories. The order of test sentences was counter-balanced for English-speaking children: half the children heard the stories in the First-Place condition first, and half heard the stories in the Second-Place condition first. However, this was found to have no effect on their answers, so the order of test sentences was fixed for the Mandarin-speaking children. These children all heard the stories in the First-Place conditions first, followed by the ones in the Second-Place condition.

The children were tested individually in a quiet corner of their daycare centre or kindergarten. The task was preceded by warm-up trials in which the puppet made several statements about a story which were obviously true or obviously false. This let the children know the puppet could say something wrong and familiarised them with the task.

4. Results

We coded each subject's initial response to the test sentences. Self-corrections were recorded only if the test sentence had not been repeated. If children changed their answer after the test sentence was repeated, this was coded as a 'mis-match' answer.

4.1 English Results

Seven English-speaking children were excluded from the final analysis because they failed to respond correctly to more than one filler item (2 children), or failed the before-and control (2 children), or they gave a mismatched answer to this control item (3 children). The remaining 17 children ranged in age from 3;4 to 5;1 (9 girls, 8 boys, mean age 4;4).

The English-speaking children accepted First-Place condition stories 91% of the time (31/34 trials), and rejected Second-Place condition stories 88% of the time (30/34 trials). The children's justifications for their rejections typically showed they understood the test sentences as meaning that the referent of the subject NP had come first. For example, one child aged 4;4 responded as follows to the Second-Place condition test sentence *The giraffe found his ball before Winnie-the-Pooh or the mouse*:

(12)   Child: no
       Puppet: no? can you help me?
       Child: the mouse found his ball first

Anna Notley, Peng Zhou, Britta Jensen, Stephen Crain

The 20 English-speaking adults tested accepted the test sentences in the First-Place condition 100% of the time (40/40 trials) and rejected the test sentences in the Second-place condition 97.5% of the time (39/40 trials).

4.2 Mandarin Results

All 20 Mandarin-speaking children (who were slightly older than the English-speaking children) successfully passed all fillers and controls. The Mandarin-speaking children accepted the test sentences in the First-Place condition 100% of the time (40/40 trials), and rejected the test sentences in the Second-Place condition 70% of the time (28/40 trials). Further examination of the data revealed that older children produced a different pattern of responses to younger children. We therefore divided the children into two age groups, a younger group of 14 children (4;6-4;7) and an older group of 6 children (5;0-5;4). The 14 younger Mandarin-speaking children accepted the test sentences in the First-Place condition stories 100% of the time (28/28 trials), and rejected them in the Second-Place condition 100% of the time (28/28 trials). The children's justifications for their rejections typically showed they understood the test sentences as meaning that the referent of the subject NP had come first. For example, one child aged 4;6 responded as follows to the Second-Place condition test sentence *Wugui zai yu huozhe xiaoma zhiqian nadao-le beike* ('The turtle got his shell before the fish or the horse'):

(13) Puppet: *Wo shuodui-le ma?*
      'Am I right?'
      Child: *Budui*
      'No'
      Puppet: *Weishenme?*
      'Why?'
      Child: *Yinwei yu xian nadao-le beike* [pointing to the fish]
      'Because the fish got his shell first'

The 6 older Mandarin-speaking children, on the other hand, accepted the test sentences in the First-Place condition 100% of the time (12/12 trials), and rejected them in the Second-Place condition 0% of the time (12/12 trials). This was the pattern we anticipated if children allowed disjunction to take scope over *zai…zhiqian*.

The 20 Mandarin-speaking adults tested accepted the test sentences in the First-Place condition 60% of the time (24/40 trials) and rejected them in the Second-Place condition 75% of the time (30/40 trials). These results for Mandarin-speaking adults contrast clearly with English-speaking adults. In

Mandarin, the conjunctive interpretation of disjunction is not the only reading in sentences with *zai…zhiqian.* That is, disjunction can also take scope over *zai…zhiqian* for adult speakers, making a Second-Place condition sentence true for second-place participants and a First-Place condition sentence possibly false for first-place participants (if a scalar implicature is computed). Indeed, the Mandarin-speaking adults who accepted the test sentences in the Second-Place condition, rejected them in the First-Place condition because they felt the puppet should have used a conjunctive statement rather than a disjunctive one. The older group of children behaved much like these adults in that they judged the Second-Place test sentences to be true. However, these children also accepted the First-Place test sentences. In fact, this is not surprising, as it has been shown that children are less likely than adults to compute scalar implicatures, especially in certain tasks like the truth value judgement task (Gualmini *et. al*. 2001, Guasti *et. al.* 2005). It is thought this is not because children lack the notion of information strength, but because they lack the computational resources needed to mentally construct an alternative representation of the sentence under consideration and then compare the relative information strength of this alternative sentence to the test sentence (Gualmini *et. al.* 2001).

4.3 Comparing the English and Mandarin Results

The comparison of the English-speaking and younger Mandarin-speaking children's and adults' results across languages is given in Figure 1.

The crucial finding was that the English-speaking children and the younger Mandarin-speaking children overwhelmingly accepted First-Place condition sentences (91% of the time in English, 100% of the time in Mandarin), and rejected Second-Place condition sentences (88% of the time in English, 100% of the time in Mandarin). This shows that both groups of children were computing a conjunctive interpretation for disjunction in the scope of *before* and *zai…zhiqian* respectively. Strikingly, younger Mandarin-speaking children's responses were more like the responses of English-speaking children and adults than like those of Mandarin-speaking adults. At age 5, Mandarin-speaking children begin to adopt more adult-like interpretations of the sentences tested.
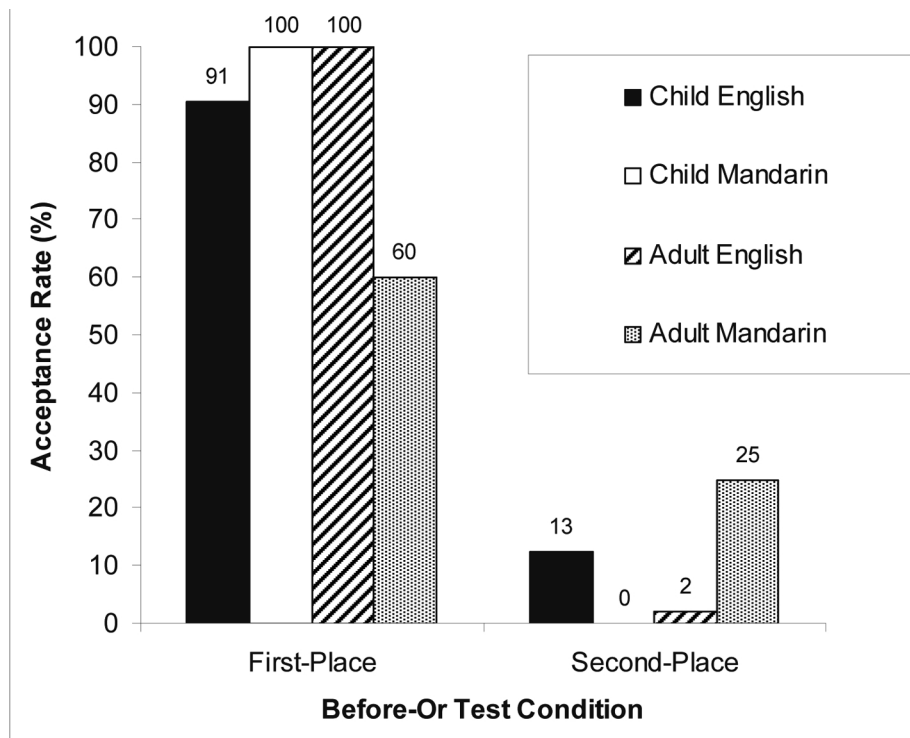
Anna Notley, Peng Zhou, Britta Jensen, Stephen Crain

Figure 1: Child and Adult Acceptance Rates Across Languages in the Before-Or Truth Value Judgement Task

5. Discussion

This study asked how children interpret disjunction in universally quantified downward entailing environments. In particular, we were interested in whether the observed cross-linguistic patterns of interpretation of disjunction in negatively flavoured DE environments could also be found in universally quantified DE environments. We identified a cross-linguistic difference in how disjunction is interpreted in sentences with *before* in English and with *zai…zhiqian* in Mandarin Chinese. In English a conjunctive interpretation of disjunction arises in the scope of *before*. We suggested that this interpretation is triggered by the presence of the universal quantifier in the semantics of *before,* which is responsible for a logical set relation process driving the conjunctive interpretation of disjunction. In a sentence of type 'A before B', disjunction in English splits event B into two 'subsets'. For event A to have occurred before

event B, it must have occurred before every time point in event B – and therefore before every time point in both subset events. In Mandarin Chinese, on the other hand, disjunction can take scope over *zai…zhiqian*. This means the logical relation between the subsets of event B and event A no longer holds.

Given these differences between English and Mandarin, we pointed out that if children rely on the scope relations attested in their language, then we expected English-speaking children to compute the conjunctive interpretation of disjunction in sentences with *before*, while Mandarin-speaking children might show a mix of both the conjunctive and non-conjunctive interpretation of disjunction in sentences with *zai…zhiqian*. By contrast, if children adhere to logical relations dictated by the hidden universal quantifier in the semantics of *before*, then we expected children across languages to initially compute the conjunctive interpretation of disjunction in sentences with *before* across languages, regardless of the scope relations attested by adult speakers of the target language.

Our data firmly support the hypothesis that children are guided by logical principles governing the interpretation of disjunction in the scope of downward entailing operators such as *before*. As we have seen, both English-speaking and younger Mandarin-speaking children clearly and consistently interpret sentences like *The dog reached the finish line before the turtle or the bunny* to mean that the dog reached the finish line first (before the turtle <u>and</u> before the bunny). They accepted such sentences in First-Place condition stories, and they rejected such sentences in Second-Place condition stories. They normally corrected the puppet in Second-Place condition stories by pointing out who really had come first. This behaviour was in line with how English-speaking adults interpret such sentences, but was quite different to how Mandarin-speaking adults interpret such sentences. When Mandarin-speaking children reach age 5 they begin to allow disjunction to take scope over *zai…zhiqian* like Mandarin adults.

The fact that Mandarin-speaking children do not initially interpret sentences containing *zai…zhiqian* and disjunction in the same way as Mandarin-speaking adults is difficult to explain on an input-based model of acquisition. In fact, an input-based model in this area would have trouble accounting for the fact that children so consistently assign any interpretation at all to these sentences at such a young age, as the type of construction under investigation rarely appears in the input to children.

This work extends previous work in the domain of children's interpretation of disjunction in the scope of negatively-flavoured downward entailing environments to a wider cross-section of downward entailing operators. Even in universally quantified downward entailing linguistic environments, children

appear to adhere to logical principles. This suggests that there may exist a deep semantic relationship between disjunction and downward entailing operators in general, and that children exploit this relationship as a linguistic universal during the language acquisition process. Further child acquisition evidence from languages in which disjunction can take scope over *before* (*e.g.,* Hungarian, Japanese or Danish) will serve to clarify this hypothesis.

References

Anscombe, G.E.M. 1964. Before and after. *The Philosophical Review* 73:3-24.

Boster, Carole T., and Crain, Stephen. 1993. On children's understanding of Every and Or. In *Conference Proceedings: Early Cognition and the Transition to Language*. Austin: University of Texas

Crain, Stephen, and Thornton, Rosalind. 1998. *Investigations in Universal Grammar: A guide to experiments on the acquisition of syntax and semantics*. Boston: Massachusetts Institute of Technology.

Crain, Stephen, Gardner, Amanda, Gualmini, Andrea, and Rabbin, Beth. 2002. Children's command of negation. In *Proceedings of the 3rd Tokyo Conference on Psycholinguistics*, ed. Yukio Otsu, 71-95. Tokyo: Hituzi Publishing Company.

Crain, Stephen, Gualmini, Andrea, and Pietroski, Paul. 2005. Brass tacks in linguistic theory: Innate grammatical principles. In *The Innate Mind: Structure and Contents*, eds. Peter Carruthers, Stephen Laurence and Stephen Stich, 175-197. Oxford: Oxford University Press.

Crain, Stephen, Goro, Takuya, and Thornton, Rosalind. 2006. Language acquisition is language change [January]. *Journal of Psycholinguistic Research* 35:31-29.

Crain, Stephen, and Thornton, Rosalind. 2006. Acquisition of syntax and semantics. In *Handbook of Psycholinguistics*, eds. Matthew J. Traxler and Morton A. Gernsbacher, 1073-1110. Amsterdam: Academic Press, Elsevier Inc.

Goro, Takuya, and Akiba, Sachie. 2004a. The acquisition of disjunction and positive polarity in Japanese. In *WCCFL 23 Proceedings*, eds. V. Chand, A. Kelleher, A.J. Rodriguez and B. Schmeiser, 251-264. Somerville, MA: Cascadilla Press.

Goro, Takuya, and Akiba, Sachie. 2004b. Japanese disjunction and the acquisition of positive polarity. In *Proceedings of the 5th Tokyo Conference on Psycholinguistics*, ed. Yukio Otsu, 137-162. Tokyo: Hituzi Shobo Publishing Company.

Grice, Paul H. 1975. Logic and Conversation. In *Syntax and Semantics III: Speech Acts*, eds. Peter Cole and Jerry L. Morgan. New York: Academic Press.

Gualmini, Andrea, Crain, Stephen, Meroni, Luisa, Chierchia, Gennaro, and Guasti, Maria Teresa. 2001. At the semantics/pragmatics interface in child language. Paper presented at *SALT XI: Proceedings from Semantics and Linguistics Theory XI*, Cornell University, Ithaca, NY.

Gualmini, Andrea, and Crain, Stephen. 2002. Why no child or adult must learn de Morgan's laws. In *BUCLD 26 Proceedings*, eds. Barbora Skarabela, Sarah Fish and Anna H.-J. Do, 243-254. Somerville MA: Cascadilla Press.

Gualmini, Andrea, Meroni, Luisa, and Crain, Stephen. 2003. An asymmetric universal in child language. In *Proceedings of the Conference "sub7 - Sinn und Bedeutung"*, ed. Matthias Weisgerber. Konstanz, Germany: Fachbereich Sprachwissenschaft der Universitat Konstanz.

Gualmini, Andrea. 2005. *The ups and downs of child language: Experimental studies on children's knowledge of entailment relationships and polarity phenomena*: Outstanding Dissertations in Linguistics. New York, London: Routledge.

Gualmini, Andrea, and Crain, Stephen. 2005. The structure of children's linguistic knowledge. *Linguistic Inquiry* 36:463-474.

Guasti, Maria Teresa, Chierchia, Gennaro, Crain, Stephen, Foppolo, Francesca, Gualmini, Andrea, and Meroni, Luisa. 2005. Why children and adults sometimes (but not always) compute implicatures. *Language and Cognitive Processes* 20:667-696.

Heinamaki, Orvokki. 1972. Before. *Chicago Linguistic Society* 8:139-151.

Horn, Laurence R. 1996. Presupposition and implicature. In *The Handbook of Contemporary Semantic Theory*, ed. Shalom Lappin, 299-319. Oxford, Cambridge, MA: Blackwell Publishers Inc.

Su, Esther, and Crain, Stephen. forthcoming. Disjunction and universal quantification in child Mandarin. In *The Proceedings of the Tenth Tokyo Conference on Psycholinguistics*, ed. Yukio Otsu. Tokyo, Japan: Hituzi Syobo Publishing.

Szabolcsi, Anna. 2002. Hungarian disjunctions and positive polarity. In *Approaches to Hungarian 8*, eds. Kenesei and Siptar, 217-241. Budapest: Akademiai Kiado

Stepwise Loss of Verb Movement[1]

Kristine Bentzen and Þorbjörg Hróarsdóttir
University of Tromsø, Norway

Abstract

This paper addresses the loss of the verb movement traditionally labelled V-to-I. Based on synchronic data it is shown that this type of verb movement corresponds to two different patterns, Long and Short non-V2 verb movement, and we propose that the change from a grammar with Subject–Verb inversion to a grammar without verb movement may proceed in a stepwise fashion, involving both these two non-V2 verb movements. This proposal is discussed in relation to loss of verb movement in English, the Romance languages French, Italian and Spanish, the Mainland Scandinavian languages, and the insular Scandinavian language Faroese.

1. Introduction

In this paper we discuss the loss of verb movement in a cross-linguistic perspective. Several people have studied the loss of Verb Second (V2) in languages like English (see *e.g.,* van Kemenade 1987, Roberts 1993; 1996, Lightfoot 1999) and French (Adams 1987a,b, Vance 1997, Roberts 1993). However, V2 has been argued not to be a single phenomenon. Rather, it has been suggested that V2 involves several micro V2 parameters, triggering verb movement to various positions in the CP-domain (*cf.* Westergaard and Vangsnes 2005). Consequently, if V2 is not one single phenomenon, the loss of Subject-Verb inversion potentially does not proceed as one single change either, but rather in a step-wise fashion, affecting various types of V2 at various stages (cf. Warner 2007, Westergaard 2009).

    In recent work, we have argued that 'V-to-I' also is not a single phenomenon (cf. Bentzen 2007, Wiklund *et al.* 2007). Based on data from Scandinavian languages, we have suggested that there are at least two types of verb movement corresponding to what is usually labelled 'V-to-I'. In this paper,

---

Kristine Bentzen and Þorbjörg Hróarsdóttir

we will explore the loss of verb movement in English, Romance and Scandinavian, focussing especially on the 'V-to-I' types of verb movements (and leaving the various types of Subject-Verb inversion (V2) aside). More specifically, we propose that the change from a grammar with (any kind of) Subject-Verb inversion to a grammar without verb movement may proceed in a step-wise fashion, and that there are two potential intermediate steps in the change from a V2/S–V inversion grammar to a grammar without verb movement:

(1) **Step I:** V2/Subject–Verb inversion
    Step II: Long non-V2 verb movement
    Step III: Short non-V2 verb movement
    Step IV: No verb movement

We support this proposal by showing that various languages have gone through some or all of these stages in their loss of verb movement over the centuries. The paper is structured as follows. In section 2 we provide some background on the two 'V-to-I' types of verb movement we have argued for on the basis of Scandinavian varieties. In section 3, we outline the history of the loss of verb movement in English main clauses. In section 4, we look at the loss of verb movement in Romance languages, still focussing on main clauses, and in section 5, we discuss the loss of verb movement in embedded clauses in Scandinavian languages. Finally, in section 6, we propose some ideas on what may lead to the loss of the various types of 'V-to-I' movement.

2. Background on the types of 'V-to-I'

The Mainland Scandinavian languages (MSc) Swedish, Norwegian and Danish are asymmetric V2 languages, with S–V inversion in main clauses, but generally no verb movement in embedded clauses. In contrast, Icelandic displays some verb movement in embedded contexts as well, and this verb movement is independent of V2. Characteristics of this non-V2 verb movement are that it obligatorily crosses negation and adverbs, (2a-b), and that in clauses with multiple adverbs, the verb has to cross all the adverbs, (2c) (examples from Wiklund et al. 2007):

(2) a. …að Jón {**keypti**} ekki {***keypti**} bókina.   (Icelandic)
        that John bought not   bought book.the
      '… that John didn't buy the book.'

b. Ég veit [af hverju Heiða {**kaupir**} oft {***kaupir**} skó.
    I   know why   Heiða buys   often buys   shoes
    'I know why Hedda often buys shoes.'

c. …first einhverjir stúdentar {**skiluðu**} sennilega {***skiluðu**}   oft
    as   some   students handed.in probably   handed.in often
    verkernum.
    assignments
    '… as some students probably handed in assignments often.'

We label this type of movement *Long non-V2 verb movement* because it targets a position higher than all adverbs in the clause.

However, there is also another type of non-V2 verb movement in Scandinavian. Although MSc generally does not show any verb movement in embedded clauses (outside of V2 contexts), there are certain dialects that do allow some non-V2 verb movement. This is illustrated with examples from Regional Northern Norwegian (ReNN)[2] in (3) (see Bentzen 2007, Wiklund et al 2007). This verb movement is different from that in Icelandic in certain respects. It strictly cannot cross negation, (3a), but optionally crosses adverbs (3b,c), and in clauses with multiple adverbs, the verb may intervene between the adverbs, (3c).

(3) a. … at   han Jon {***kjøpte**} ikke {**kjøpte**} boka.   (ReNN)
        that he John bought not   bought book.the
        '… that John didn't buy the book.'

   b. Æ vet [koffer ho Hedda {**kjøpe**} ofte {**kjøpe**} sko].
        I   know why   she Hedda buys   often buys   shoes
        'I know why Hedda often buys shoes.'

   c. … ettersom nån   studenta {**leverte**} sannsynligvis {**leverte**} ofte
        as   some students handed.in probably   handed.in often
        oppgava.
        assig'nts
        '… as some students probably handed in assignments often.'

We label this type of movement *Short non-V2 verb movement* because it appears to have landing sites at various positions (both high and low) in the IP domain, as seen by its relative position with regard to adverbs, especially in cases of multiple adverbs.[3]

---

[2] ReNN refers to the dialects spoken in the more rural parts of Northern Norway.

[3] Of course, we are aware of the fact that it may be problematic to use the position of the negation

Kristine Bentzen and Þorbjörg Hróarsdóttir

Along with Wiklund *et. al.* (2007) we take these two types of non-V2 verb movement to target different domains in the clause; the Long non-V2 verb movement targets the lowest projection of the Rizzian CP domain, FinP (cf. Rizzi 1997), whereas the Short non-V2 verb movement targets a functional projection in the IP domain. Furthermore, we follow Bentzen (to appear a) in assuming different triggers for the two verb movement types. Bentzen proposes that the Long non-V2 verb movement is triggered by a Finiteness feature on the Fin head. There is parametric variation with respect to how the [Fin(iteness)] feature is licensed, either through Move or through Agree. Languages that license [Fin] through Move display overt verb movement to FinP, that is, Long non-V2 verb movement. Languages that license [Fin] through Agree, on the other hand, will not show Long non-V2 verb movement.[4] The Short non-V2 verb movement Bentzen (to appear a,b) links to predication licensing (or the EPP, if you like). The feature [Pred(ication)] may be associated with various projections in the IP domain, and is licensed by overt movement to the specifier of the head carrying this feature. Again, there may be parametric variation concerning how [Pred] is licensed. Either the subject moves alone to the specifier of the head carrying [Pred] through so-called spec-raising, or the subject piedpipes the vP in moving to this projection, through so-called spec-piedpiping. (See Biberauer and Richards 2006 for a similar proposal, and for the coinage of the terms *spec-raising* and *spec-piedpiping*). Languages that allow the spec-piedpiping option will display the effects of Short non-V2 verb movement, whereas languages that may only license [Pred] through spec-raising will not have Short non-V2 verb movement.

We take these two patterns of non-V2 verb movement as a starting point and ask whether we see instances of such patterns in the diachronic loss of verb movement. More specifically, we explore the hypothesis that the loss of verb movement may proceed in a step-wise fashion, going through the two potential intermediate steps on the way to a grammar with no verb movement.

3. Loss of verb movement in English

In present day English finite auxiliaries of course perform both S–V inversion

---

as a diagnostic for verb movement cross-linguistically, as the position of negation is not necessarily the same from one language to another. We therefore investigate the verb's placement relative to both negation and adverbs in the various languages.

[4] Like Wiklund et al. 2007 and Bentzen to appear a, we take the Long non-V2 verb movement to be phrasal movement of vP to SpecFinP, but nothing in the current paper hinges on this assumption.

and move across negation and adverbs in main clauses, whereas finite main verbs do not show this behaviour. However, at earlier stages of English, also main verbs displayed these characteristics. In the history of English, we find changes with respect to the placement of finite main verbs involving Steps I, II and IV.

3.1 Old English

Old English displayed S–V inversion in various contexts where such word order is not found today (e.g. van Kemenade 1987, Roberts 1993, EyÞórsson 1996, Warner 2007). After certain introductory elements, like *Þa, Þonne* and *wh*-words, the V-S order is attested with both auxiliaries and main verbs, and with both DP and pronominal subjects, as illustrated in (4) (from Pintzuk 1999:138 and van Kemenade 1987:138). In such clauses, the verb is argued to target a "V2-like" (high) position (cf. among others Haeberli 2002, Warner 2007).

(4) a. Þa **ge-mette** he sceaðan.   (Old English)
        then met   he robbers
        'Then he met robbers.'
    b. Hvæt **sægest** Þu   yrðling?
        what say   you-sg. ploughman
        'What do you say, ploughman?'

In clauses with other introductory elements, however, the verb precedes DP subjects, but follows pronominal subjects, as shown in (5) (examples from Haeberli 2002:245 and Pintzuk 1999:86). For such clauses it has been argued that the V–S order results from the verb moving to a lower projection (e.g. AgrP). Pronominal subjects appear in a higher specifier preceding the verb (yielding S_{Pron}–V), while DP subjects remain in a lower specifier following the verb (yielding V–S_{DP}).

(5) a. And egeslice **spæc** Gregorius be   ðam   (Old English)
        and sternly   spoke Gregorius about that
        'And Gregorius spoke sternly about that'
    b. ælc yfel he **mæg** don
        each evil he can do
        'He can do each evil'

Old English also displayed Long non-V2 verb movement. At this stage English employed a pattern involving two negative elements to express sentence negation. The negative particle *ne* always occurred to the immediate left of the

Kristine Bentzen and Þorbjörg Hróarsdóttir

finite verb, whereas the element *na* or *no* was used as the second element,[5] with the finite verb moving across this element, see (6) (example from Fischer et al. 2000: 55).

(6) Ne **bið na** se leorningcniht furðor Þonne his lareow.   (Old English)
   not is   not the apprentice   further than   his master
   'The apprentice is not ahead of his master.'

3.2 Middle English (1150-1500)

S–V inversion continued into Middle English, and verb movement across both adverbs and negation, i.e. Long non-V2 verb movement is also attested (cf. Santorini and Kroch 2007):

(7) a. Þe gong man **resortyd alwey** to Þe preste   (Middle English)
       the young man resorted always to the preist
   b. but Balyn **dyed not** tyl the mydnyghte after
       but Balyn died   not till the midnight   after
       'but Balyn did not die till the midnight after'

In the course of the Middle English period, several changes took place, which led to the Modern English grammar with respect to verb placement. In this period, the decline of verb raising across negation coincided, and probably interacted, with three further changes: (i) a simplification of the verbal agreement system, (ii) a change in the status of *not*, and (iii) the emergence of *do* support. For reasons of space we will not go into these changes and the potential interaction with the changes in English verb placement here.

   In the late 14th and early 15th centuries, S–V inversion in topic-initial constructions shows a sharp decline. By the end of the Middle English period, S–V inversion (with main verbs) was more or less lost.[6]

3.3 Early Modern English (1500-1800)

Since S–V inversion and Long non-V2 verb movement co-existed during the Old and Middle English period, one may question whether Step II as an independent stage can be attested in English. However, according to Biberauer and Roberts (to appear), there was a (brief) stage with systematic movement of both main verbs and auxiliaries across negation, in the absence of the S–V inversion option in late Middle to Early Modern English (15th-16th century). Roberts (1993) and

---

[5] In addition to *na* and *no*, the negative elements *noht* and *nawiht* are also attested at this stage.
[6] The quantitative results in the literature are rather variable here, but inversion may actually not have been completely lost until the 18th century.

Biberauer and Roberts (to appear) argue that Early Modern English displayed the word order S–V$_{Finite}$–*not* in subordinate clauses that are known to resist embedded inversion, as shown in (8) (from Roberts 1993:323):

(8) And gif he **be noght** so, then…   (Early Modern English)
   and if   he be not   so then
   'And if he's not, then…'

Hence, it seems plausible that Step II can be identified as a grammar attested in Early Modern English, where Long non-V2 verb movement was employed independently of S–V inversion (Step I).

3.4 Modern English

In present day English, finite main verbs do not undergo S–V inversion, nor do they move across adverbs and negation in any contexts. In for examples *wh*-questions and clauses with negation, *do*-insertion is employed instead, as in (9a-b), whereas in clauses with adverbs the finite main verb simply remains in situ, (9c).

(9) a. Where **did** he go?   (Modern English)
   b. I **did** not see him.
   c. We never **met** her.

     Summing up the changes found in English, we thus find that Step I (S–V inversion) was attested in Old and Middle English; Step II (Long non-V2 verb movement) was attested in Old, Middle and Early Modern English, and Step IV (no verb movement) corresponds to present day English.

4. Loss of verb movement in the Romance languages

Early Romance varieties displayed certain characteristics found in typical V2 languages, like S–V inversion, but this is not found today in any of the three Romance languages considered here. In this section we outline changes in the placement of finite main verbs and auxiliaries in main clauses in a selection of Romance languages: French, Italian, and Spanish. Within these Romance languages, we find changes from Step I to Step II in French and from Step I to Step III (potentially coinciding with Step IV) in Italian and Spanish. We present the changes in French and the changes in Italian and Spanish separately, since the current states of these two groups are different.

4.1 Old to Middle French (approximately 900-1600)

Old French was a V2 language as illustrated by the examples in (10) (from Labelle 2007:297), showing S–V inversion in main clauses (cf. also Adams

Kristine Bentzen and Þorbjörg Hróarsdóttir

1987a,b, Roberts 1993, Vance 1997).

(10) a. Mais a la bataille n'**ose** ile pas venir.   (Old French)
   but to the fight neg dare he not come
   'but he doesn't dare come to the battle'
  b. **Fustes** vus unques baptizé ne levé?
   were you ever baptized nor raised
   'Were you ever baptized?'

Like most Modern Germanic languages, Old French was an asymmetric V2 language; that is, generally no V2 in embedded clauses, (11a) (from Adams 1987b:2). However, like in the Germanic languages, V2 was possible in the complements of bridge verbs in Old French, (11b) (from Vance 1997:141):

(11) a. Einsi corurent par mer tant que il **vindrent** à Cademelée. (Old French)
   thus ran (they) by sea unti they arrived at Cadmee
   'Thus, they travelled by sea until they arrived at Cadmee.'
  b. Et il respont que ce ne **feroit** il pas
   and he answers that this neg would-do he not
   'and he answers that he will not do this'

Old French also display instances of Long non-V2 verb movement. In embedded clauses without S–V inversion, the finite verb still precedes the negative element *pas*, as shown in (12) (from Vance 1997:134).

(12) Et il li respont qu'il ne **set** pas tres bien cui filz il fu (OF)
  and he to-her responds that-he neg knows not very well whose son he was
  'And he answers her that he doesn't know very well whose son he was.'

During the Middle French period (1300-1500) there was a gradual loss of the V2 constraint. While Old French essentially was a V2 language, Middle French constitutes a 'transitional period' (cf. Vance 1989, Roberts 1993). Hence, Middle French was an optional V2 language, as illustrated in (13) (examples from Vance 1997: 264, 268), but by the Mid-16th century V2 was more or less lost.

(13) a. Lors la royne **fist** Saintré appeller.   (Middle French)
   then the queen made Saintré to-call
   'Then the queen had Santré called'
  b. A ce cop **cognois** je bien que jamais ne vauldrez   riens.
   at this blow know I well that never neg will-be-worth-2pl nothing
   'Now that you tell me this, I am sure that you'll never be worth anything.'

4.2 Modern French

Modern French does not employ systematic S–V inversion. However, Long non-V2 verb movement is still obligatory. The claim that French has verb movement across negation deserves a note here, since the finite verb always *follows* the negative element *ne*, while it obligatorily precedes the negative element *pas*, (14a). However, *ne* is generally assumed to be cliticized onto the verb, moving along with it, hence always preceding it (cf. e.g. Zanuttini 1997, Cinque 1999). Thus, obligatory verb movement across negation here refers to the verb's position relative to *pas*. Moreover, while the verb's placement with respect to negation may not be a solid diagnostic for the position of the verb, we emphasize that verb movement across adverbs is obligatory, and crucially has to cross *all* clause-medial adverbs in the clause, (14b-c). Based on this we argue that French has Long non-V2 verb movement.

(14) a. Jean ne {**mangeait**} pas {***mangeait**}les gâteaux. (Modern French)
       Jean NEG ate   not   ate   the cakes
       'Jean didn't eat the cakes.'

   b. Jean {**embrasse**} souvent {***embrasse**} Marie.
       Jean kisses   often   kisses   Marie
       'Jean often kisses Marie.' (based on Pollock 1989:367)

   c. Marie {**mangeait**} malheureusement {***mangeait**} souvent tous les
       Marie ate   unfortunately   ate   often all the
       gâteaux.
       cakes.
       'Marie unfortunately often ate all the cakes.'

We thus see a change in the history of French from a grammar at Step I with S–V inversion to a current grammar at Step II with Long non-V2 verb movement.

4.3 Old Italian and Spanish (approximately 900-1400/1500)

Both Old Italian and Old Spanish displayed V2/S–V inversion in non-subject-initial clauses. The Italian examples in (15) are from Old Neapolitan, from early 14th century prose text (from Ledgeway 2008:440-441), and the Spanish examples in (16) are from texts from the 15th and 16th century (from Fontana 1993:65).

(15) a. maraviglyoso **era** lo suono   (Old Neapolitan)
       marvellous   was the sound
       'the sound [of the clashing swords] was tremendous'

   b. de poy queste parole **ademandao** lo   messayo   licencia
      after   these words asked   the messenger permission
      'following these words the messenger asked permission (to leave)'

(16) a. este logar **mostro** dios a abraam   (Old Spanish)
      this place showed God to Abraham
      'God showed Abraham this place.'

   b. Despues   **quiso** S.M.   ver algunas ciudades del estado de Milan
      afterwards wanted his majesty see some   cities   of state   of Milan
      'His majesty wanted to see some cities in the state of Milan afterwards.'

Old Neapolitan appears to have been an asymmetric V2 language like Old French. Old Spanish, on the other hand was a symmetric V2 language. Embedded V2 effects are found in relative clauses and subordinate clauses that are not complements of the restricted set of verbs licensing embedded V2 asymmetric V2 languages like MSc, as illustrated in (17) (cf. Fontana 1993: 72-73).

(17) Quando esto **oyo**   el Rey   (Old Spanish)
     when   this heard the king
     'When the king heard this…'

4.4 Modern Italian and Spanish

Above we argued that Modern French has Long non-V2 verb movement. The placement of verbs in Modern Italian and Spanish, however, differs from Modern French in certain respects. Again, the use of the position of negation as a diagnostic is somewhat problematic. Generally, Italian and Spanish only use one negative element *non* and *no*, corresponding to French *ne*, and the finite verb always follow these elements. However, Italian also has the negative element *mica*, which is argued to correspond to *pas* (cf. Cinque 1999, inter alia). As (18a) shows, verb movement across *mica* is not necessarily obligatory, in contrast to what we saw with *pas* in French (from Cinque 2004:704). Moreover, verb movement across adverbs is optional, (18b)-(19a), the verb does not need to cross *all* adverbs in a clause, (18c)-(19b). We thus argue that Modern Italian and Spanish display Short non-V2 verb movement.

(18) a. Di solito Maria mica **prende** il treno.   (Modern Italian)
      usually Mary not takes   the train
      'Mary usually doesn't take the train.'

   b. Alcuni studenti {**fraintendono**} spesso {**fraintendono**} il compito.
      some students misunderstood often   misunderstood the assignment
      'Some students often misunderstood the assignment.'

    c. Alcuni studenti {**fraintendono**} spesso {**fraintendono**} completamente
       some   students   misunderstood often   misunderstood completely
       il   compito.
       the assignment
       'Some students often misunderstood the assignment.'

(19) a. Algunos estudiantes {**malinterpretaron**} siempre   (Modern Spanish)
       some   students   misinterpreted   always
       {**malinterpretaron**} la tarea.
        misinterpreted   the assignment
       'Some students always misunderstood the assignment.'

    b. Algunos estudiantes {%**malinterpretaron**} a menudo
       some   students   misinterpreted   often
       {**malinterpretaron**} completamente la tarea.[7]

       misinterpreted completely the assignment
       'Some students often completely misunderstood the assignment.'

Thus, we see a change from Step I in Old Italian/Spanish to Step III in Modern Italian and Spanish. Furthermore, the fact that Short non-V2 verb movement appears to be optional in these languages, suggests that a Step IV grammar is also available in Modern Italian and Spanish.

5. Loss of verb movement in Mainland Scandinavian

As mentioned in the introduction, the Modern Mainland Scandinavian languages (MSc) are asymmetric V2 languages displaying V2 in main clauses but generally no verb movement in embedded clauses. But this has not always been the case. In the history of MSc we find changes in embedded verb placement involving Steps I, II, and IV, and in certain varieties also Step III.

5.1 Old to Middle Scandinavian (approximately 1200-1500)

Old and Middle MSc had some sort of verb movement in embedded clauses as well. There is some evidence suggesting that there was generalized embedded V2 with subject-verb inversion. This can been seen in Old Norse (example from Vikner 1995, citing Nygaard 1906:376), (20), in Old Swedish (example from Holmberg and Platzack 1995, citing Larsson 1931:75), (21), and in Middle Danish (example from Vikner 1995, citing Mikkelsen 1911:588), (22). In the present-day languages, subject-verb inversion is impossible in these types of embedded contexts.

---

[7] The "%" sign represents interspeaker variation among our Spanish informants on this point.

(20) Gjarna mundi hann hafi viljat   drepa hann í fyrstu, ef honum **væri** Þat lofat.
gladly would he   have wanted kill   him   at first   if him$_{DAT}$ were it   allowed

(21) Þa   vildi   iah slikum rætti   vnæ, sum nu   **føræ** iah, Þær fram.
then wanted I   such   redress get that now bring I   you forth
'Then I would be satisfied with such redress as I now offer to you.'

(22) Saa   som nu **giør** Jøderne low.
in-such-a-way that now make Jews.the law

Furthermore, Old and Middle MSc also displayed Long non-V2 verb movement. Verb movement across negation and adverbs is attested in non-V2 contexts as illustrated for Middle Norwegian in (23) (from Sundquist 2002:330), for Old/Middle Swedish in (24) (from Falk 1993:150 and Sundquist 2002:233), and for Middle Danish in (25) (from Hrafnbjargarson 2004:212):

(23) Jtem kiere her Eske om hann **forthewffeur** oss icke (Middle Norwegian, 1522)
also dear lord Eske if   he   waits.for   us   not
'Also, dear lord Eske, if he does not wait for us…'

(24) … hvi kristne   män **räddos** ei pino.   (Middle Swedish, 1385)
why christian men feared not pain

(25) vm min man **hafvir** inkte rætfongit   gooz.   (Middle Danish, 1425)
if   my man has   not   rightly.received goods
'if my husband has not rightly received goods'

Towards the end of this period, rich verbal morphology is lost.

5.2 The Early Modern Period (1500-1700)

By the end of the 16th century, there are very few cases of embedded V2, (apart from in certain *that*-clauses). Still we find that Long non-V2 verb movement is still dominant at the beginning of this period. This means that at this point, the Long non-V2 verb movement was an operation independent of subject-verb inversion. During the Early Modern Period from the 16th to the 18th century, the word order without verb movement in embedded clauses also appears, and the two word orders occurred side by side for a couple of centuries, as shown in (26)-(27) (from Sundquist 2003:238).

(26) om vy **for** icke de suar   (EMD, 1556-1573)
if   we get not the answers
'if we do not get the answers.'

(27) som   icke **kan** skriffuis   paa denne gang   (EMD, 1601-1625)
which not   can written-be at   this   time
'which cannot be written at this time'

Thus, Old to Middle MSc had embedded V2 and also Long non-V2 verb movement, but the former was lost by the end of the Middle MSc period, whereas the latter was gradually lost during the Early Modern Period. By 1700 Long non-V2 verb movement was lost, and the pattern with no verb movement in embedded clauses had taken over.

5.3 Modern MSc

For most Scandinavian varieties, the loss of Long non-V2 verb movement led to a pattern with No verb movement in non-V2 contexts. These varieties have thus gone through Steps I, II, and IV in the change of verb placement in embedded clauses.

(28) om vi ikke **får** svarene.   (Modern Norwegian)
    if   we not get answers.the
    'if we do not get the answers.'

However, as mentioned above, the dialect Regional Northern Norwegian (ReNN) optionally shows verb movement in non-V2 contexts that is very similar to the pattern found in modern day Italian and Spanish, namely Short non-V2 verb movement. This was illustrated in examples (2b-c). The same type of verb movement is found in these contexts in Northern Ostrobothnian Swedish (NOb), as shown in (29) (cf. Bentzen to appear c).[8] As the examples show, this dialect also disallows non-V2 verb movement across negation (29a), but optionally allows it across various adverbs (29b-c). Moreover, as (29c) shows, the verb may also intervene between multiple adverbs.

(29) a. Ja veit [fövaa   Göran {***itär**} int {**itär**} korv].   (NOb)
      I know for-what Göran   eats   not eats hot.dogs
      'I know why Göran doesn't eat hot dogs.'
    b. Ja föstoo   int [fövaa   an {**tvättar**} så tökält {**tvättar**} biln   sin].
      I understood not for-what he   cleans   so often   cleans   car.the REFL
      'I didn't understand why he cleans his car so often.'
    c. Veit du [vem an {**hade**} troligen {**hade**} tökläst kunnat be om   hjälp]?
      know you who he had   probably had   often.est could ask about help
      'Do you know who he probably had been able to ask for help most of the time?'

Presumably, both ReNN and NOb required Long non-V2 verb movement at an

---

[8] Northern Ostrobothnian Swedish is spoken in the Ostrobothnia area in Finland, and includes the variety discussed in the literature as the Kronoby dialect. The findings reported on in Bentzen to appear c thus confirm Anders Holmberg's observations on verb movement in the Kronoby dialect in the late 1980s (see Platzack & Holmberg 1989).

earlier stage, just like older varieties of Norwegian and Swedish in general did. Therefore, one might argue that these dialects have gone through the change from a Step II grammar to a grammar where Step III and Step IV cooccur. However, there is a Scandinavian language in which all the four steps are more convincingly instantiated, namely the insular Scandinavian language Faroese. We turn to that in the next section.

5.4 Faroese: A showcase for step-wise loss of verb movement

Faroese is a language which descends from Old Norse in the 9th and 10th century (when people from Norway settled on the Faroe Islands). It thus shares its ancestor language with Norwegian. As we have already seen, Old Norse displayed some sort of V2/S–V inversion in what we refer to as non-V2 contexts (cf. example (20)). Examples from the oldest authentic Faroese written document, the *Seyðabræv* ('Sheep Documents'), dating back to 1298, also suggest that Old Faroese had obligatory Long non-V2 verb movement in non-V2 contexts. The word order V-Neg is found consistently in embedded (non-V2) clauses (examples from Thráinsson et al. 2004).

(30) a. Enn ef han **uill**   æigi læigu taca   (Old Faroese, 1298)
        but if he wants not rent   take
        'But if he doesn't want to charge rent...'
    b. Þær   nauðsyniar hava til gengit at   han **matti** æigi or   fora
        those causes   have occurred that he   could not from move
        'that such important things have occurred that he could not remove [the sheep]'
    c. ok onnur elld gogn   er hon **mintist**   ei huorsu morg voru
        and other fire utensils that she remembered not how   many were
        'and other cooking utensils that she didn't remember how many they were'

Around the 19th-20th century, Long non-V2 verb movement appears to have become optional in Faroese. This is evident both in the 1832 translation of St. Matthew, and in Jakob Jakobsen's 'translation' of the Seyðabræv (from 1907) into 19th century Faroese. The examples in (31)-(32) show that in both these two documents, Long non-V2 verb movement and lack of verb movement cooccur, seemingly also within individual speakers/writers. The examples are from Thráinsson *et al.* (2004).

(31) a. á Moudstandarin **skeâl** ikkje antvora   Doumarinun té (Faroese, 1832)
        so that opponent.the   shall not   hand-over judge.the   you
        'so that the opponent shall not hand you over to the judge'

   b. O   um Neâgar ikkje **víl**   teâga uimoudi Tikun
       and if anybody not wants take with   you
       'and if anybody does not want to accept you'

(32) a. Men um hin   ikki **vil**   taka leigu…   (Faroese, 1907)
       but if the-other not wants take rent

   b. sum gjørdu, at hann **kundi** ikki føra seyð sín burtur
      that caused that he   could not move sheep REFL *away*
   c. og onnur kókingaramboð, ið hon **mintist ikki** kvussu mong ið   vóru
      and other cooking-utensils that she remembered not how   many that were

This optionality has lasted up until recently, at least until the Mid 20th century (see Petersen 2000, Thráinsson 2003; examples are from Jonas 2002).

(33) Han spyr, hví tað {**eru**} ikki {**eru**} fleiri tílíkar samkomur. (Faroese, Mid 1900)
    he asks why there are not are   more such gatherings

(34) Hóast   fólk   {**hafa**} ongantíð {**hafa**} fingið fisk her…
    although people have never   have   caught fish here

Recently, however, it has been argued in several studies that the Long non-V2 verb movement is no longer a characteristic feature in Faroese. Rather, results from recent fieldwork in the Faroe Islands reported on in Bentzen et al. (to appear) show that modern day Faroese display a pattern that resembles ReNN and NOb (see also Thráinsson 2003, Heycock and Sorace to appear). Like the dialects of MSc, Faroese today hardly allows verb movement across negation, (35a), but verb movement across and in between adverbs is still optionally available for most speakers, (35b-c).

(35) a. Hon fortaldi mær [hví Ása {***etur**} ikki {**etur**} blóðpylsu]. (Mod. Faroese)
      she told   me why Ása   eats not   eats blood.saussage
      'She told me why Ása doesn't eat blood saussages.'
   b. Jógvan skilir ikke [hví Kjartan {**vaskar**} so ofta {**vaskar**} hansara bil].
      Jógvan understands not why Kjartan cleans so often cleans his car
      'Jógvan doesn't understand why Kjartan cleans his car so often.'
   c. Hevur tú hoyrt [hvørjum hann {**fer**} allarhelst {**fer**} sum oftast í biograf
      have you heard who he goes probably goes as often-est in cinema
      saman víð]?
      together with
    'Did you hear who he probably most of the time goes to the cinema together with?'

As the data presented above suggest, Faroese has indeed gone through all four steps in the loss of (embedded) verb movement: Step I attested in Old Norse, Step II attested in Old Faroese up until the Mid 20th century, and Steps III and IV currently attested, with Step III already being lost for some speakers.

6. What triggers the loss of non-V2 verb movement?

In this section we will offer some brief speculations on what may lead to the loss of Long and Short non-V2 verb movement. Recall from section 2 that we take Long non-V2 verb movement to target FinP and to be triggered by [Fin(iteness)] licensing, whereas Short non-V2 verb movement was linked to the licensing of a flexible [Pred(ication)] feature which may be associated with various projections in the IP domain.

Traditionally, the loss of V-to-I movement is correlated with the loss of rich verbal agreement morphology (cf. among many others Vikner 1995 and Bobaljik and Thráinsson 1998). The type of verb movement discussed in the literature typically involves movement across negation, and we therefore take the proposed link to hold between rich agreement morphology and what we call Long non-V2 verb movement. More recently, Biberauer and Roberts (2005, 2008, to appear) have proposed that verb movement across negation was lost in English because of a combination of the loss of V2/S–V inversion and *tense* morphology becoming poorer. They argue that the loss of S–V inversion yields an SVO order which is ambiguous between Long (or Short, for that matter) non-V2 verb movement or no verb movement. As a consequence, the trigger for verb movement becomes less robust, as most declarative clauses (without negation or adverbs) will not provide any cues about the position of the verb. However, according to Biberauer and Roberts, rich tense marking also provides a cue for verb movement across negation, so a language with sufficiently rich tense but without V2 may still display what we label Long non-V2 verb movement.

This analysis works well for English. S–V inversion was lost during the Middle English period, and as tense marking was poor in English by that stage, (independent) Long non-V2 verb movement was a short-lived phenomenon in the late Middle to Early Modern English period. Their analysis also seems plausible for the loss of Long non-V2 verb movement in MSc embedded clauses. Embedded V2 was lost during Middle Scandinavian, and 'fairly rich' tense morphology declines in this period and is finally lost in the first part of the Early Modern period in MSc (Bandle 2005). Long non-V2 verb movement declines during the Early Modern period, so it seems plausible that independent Long non-V2 verb movement in the absence of rich tense was a short-lived

phenomenon also in MSc.

However, recall that Faroese has displayed embedded independent Long non-V2 up until at least the Mid 20th century. Under the approach by Biberauer and Roberts that would be surprising given that embedded V2 does not seem to be attested after the Old Faroese period. Not much is known about the diachronic perspective of verbal morphology in Faroese, but from the data available, verbal morphology (both tense and agreement) appears to be 'fairly poor' at a stage centuries before the 1950s. This raises the question of what triggered Long non-V2 verb movement up until it was lost in the Mid 20th century. Another case that questions the correlation between rich morphology and Long non-V2 verb movement is the synchronic situation in French. French is a non-V2 language, and in the spoken language, only about 4-6 tenses are distinguished and agreement morphology is in fact quite poor in actual pronunciation. Still Long non-V2 verb movement is a consistent feature of this language. Moreover, modern Italian and Spanish distinguish at least 5-7 (or more) tenses and have rich agreement morphology, but here Long non-V2 verb movement is not an available. It is therefore not clear whether the loss of V2/S–V inversion combined with the loss of rich tense (or agreement) morphology is able to provide the correct analysis of why Long non-V2 verb movement is lost in a language. It should also be noted that these paradigmatic approaches to whether a language has verb movement or not are problematic from an acquisition point of view. In child language, verb placement is generally in place at quite an early age (cf. among many others Westergaard 2005), and often before the fully-fledged verbal paradigm is in use in the child's language. Thus, although it is possible that the loss of V2/S–V inversion and the loss of rich verbal morphology contributed to the loss of Long non-V2 verb movement in English and in MSc, this cannot account for all the languages we have looked at here. Another potential analysis is that the loss of this type of verb movement is linked to a reanalysis of negation from a phrase to a head. For reasons of space, we will not go into that possibility here.

With respect to the loss of Short non-V2 verb movement, we again take an analysis of English proposed by Biberauer and Roberts (2005) as the starting point. Biberauer and Roberts analyse verb movement in Old and Middle English as vP raising. This thus resembles our approach to Short non-V2 verb movement as vP piedpiping. Furthermore, Biberauer and Roberts suggest that there is a correlation between the decline of vP raising and the emergence of clause-internal expletives in English. Clause-internal expletives were optionally available from the 13th century as a way of satisfying T's EPP feature alongside vP raising. At the beginning of the Early Modern English period after the loss of

OV, the new SVO order was ambiguous with respect to verb movement, and vP raising was, according to Biberauer and Roberts, alternatively reanalysed as subject movement. This led to an increased use of clause-internal expletives in the position previously occupied by vP, and further decreased the robustness of the trigger for vP raising. Eventually, DP-raising and expletive insertion became the only ways of satisfying T's EPP feature, that is, of licensing predication in our terms.

We potentially find a similar diachronic correlation between the emergence of a clause-internal expletive and the loss of verb movement in MSc. The 'new' word order Adv-V appeared from the 16th century onwards, and it became obligatory by 1700. During the 15th century, expletives appeared in MSc, first clause-initially, and later clause-internally, and they became obligatory in the clause-internal position in the course of the next couple of centuries. Thus, it appears that the loss of the V-Adv word order and the emergence of obligatory clause-internal expletives coincide in the history of MSc. Note that it is of course difficult to know whether the emergence of expletives is linked to the loss of Long or Short non-V2 verb movement. All diachronic research on this topic (as far as we know) has treated verb movement across negation and across adverbs as one single phenomenon, and most of the data provided in the literature involves verb movement across negation. However, there are some synchronic correlations suggesting that it is the Short non-V2 verb movement that is relevant to the presence or absence of clause-internal expletives. Italian and Spanish are pro-drop languages and do not have expletives at all, and in these languages, (at least some) Short non-V2 verb movement appears to be obligatory; the verb cannot follow very low adverbs like *completely* (cf. e.g. Cinque 1999). In contrast, in standard modern MSc, clause-internal expletives are obligatory, and Short non-V2 verb movement is not available. Finally, in modern Faroese, clause-internal expletives are optional, and so is Short non-V2 verb movement. These observations support the proposed diachronical link between clause-internal expletives and Short non-V2 verb movement.

7. Summary and Concluding Remarks

In this paper we have argued that there are (at least) two intermediate verb movement grammars between a grammar with (any kind of) S–V inversion and a grammar without verb movement, viz. Long and Short non-V2 verb movement. We have provided both diachronic and synchronic data to support this argument. Furthermore, we have suggested that the loss of the verb movement traditionally labelled V-to-I may occur in a step-wise fashion involving these two intermediate

stages. Faroese was given as the prime example of a language that apparently has gone through all four stages in the loss of verb movement in embedded clauses.

In the final section we addressed what may trigger the loss of the non-V2 verb movements. We discussed the possible link between the loss of Long non-V2 verb movement and the loss of V2 combined with the loss of rich verbal morphology. As pointed out, there are various challenges to this proposal. The loss of Short non-V2 verb movement was linked to the emergence of clause-internal expletives, and this proposal received both diachronic and synchronic support. There are of course still several open questions about what triggers the loss of various types of verb movement, but we will have to leave for future research.

References

Adams, Marianne. 1987a. *Old French, Null Subjects, and Verb Second Phenomena*. PhD dissertation, UCLA.

Adams, Marianne. 1987b. From Old French to the theory of pro-drop. *Natural Language and Linguistic Theory* 5 1:1–23.

Bandle, Oskar. 2005. *The Nordic Languages. An International Handbook of the History of the North Germanic Languages*. Berlin/New York: Mouton de Gruyter.

Bentzen, Kristine, Piotr Garbacz, Caroline Heycock, and Gunnar Hrafn Hrafnbjargarson. to appear. On variation in Faroese verb placement. In *Nordlyd 36: Faroese Syntax 2009*, ed. Peter Svenonius, Kristine Bentzen, Janne Bondi Johannessen, Caroline Heycock, Jógvan i Lon Jacobson, Jeffrey Parrott, Tania Strahan, and Øystein Alexander Vangsnes. *Tromsø Working Papers in Language and Linguistics*.

Bentzen, Kristine. 2007. *Order and Structure in Embedded Clauses in Northern Norwegian*. PhD dissertation, University of Tromsø, Norway.

Bentzen, Kristine. To appear a. Splitting up the V-to-I parameter. *NELS 39 Proceedings*.

Bentzen, Kristine. To appear b. Subject positions and their interaction with verb movement. *Studia Linguistica*.

Bentzen, Kristine. To appear c. Kronoby revisited: Verb movement in embedded non-V2 contexts in Northern Ostrobothnian. In *Nordlyd 35: Solf Proceedings*, ed. Øystein Alexander Vangsnes, Gunnar Hrafn Hrafnbjargarson, and Christine Bjerkan Østbø. *Tromsø Working Papers in Language and Linguistics*.

Biberauer, Theresa and Ian Roberts. 2005. Changing the EPP parameters in the history of English: Accounting for variation and change. *English Language and Linguistics* 9:5–46.

Biberauer, Theresa and Ian Roberts. 2008. Cascading parameter changes: Internally-driven change in Middle and Early Modern English. In *Grammatical Change and Linguistic Theory*, ed. Þórhallur Eyþórsson, 79–113. Amsterdam/Philadelphia: John Benjamins. The Rosendal Papers.

Kristine Bentzen and Þorbjörg Hróarsdóttir

Biberauer, Theresa and Ian Roberts. To appear. Subjects, tense and verb movement. In *Parametric Variation: Null Subjects in Minimalist Theory,* ed. Ian Roberts and Anders Holmberg. Cambridge: Cambridge University Press.

Biberauer, Theresa and Marc Richards. 2006. True optionality: When the grammar doesn't mind. In *Minimalist Essays*, ed. Cedric Boeckx, 35–67. Amsterdam: John Benjamins.

Bobaljik, Jonathan David, and Höskuldur Thráinsson. 1998. Two heads aren't always better than one. *Syntax* 1.1: 37-71.

Cinque, Guglielmo. 1999. *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. New York: Oxford University Press.

Cinque, Guglielmo. 2004. Issues in adverbial syntax. *Lingua* 114:683–710.

EyÞórsson, Þórhallur. 1996. Functional categories, cliticization, and verb movement in the early Germanic languages. In *Studies in Comparative Germanic Syntax II*, eds. Höskuldur Thráinsson, Samuel David Epstein, and Steve Peter, 109–139. Dordrecht: Kluwer.

Falk, Cecilia. 1993. Non-referential subjects and agreement in the history of Swedish. *Lingua* 89:143–180.

Fischer, Olga, Ans van Kemenade, Willem Koopman, and Wim varn der Wurff. 2000. *The Syntax of Early English*. Cambridge: Cambridge University Press.

Fontana, Josep M. 1993. *Phrase Structure and the Syntax of Clitics in the History of Spanish*. Ph.D. thesis, University of Pennsylvania, Philadelphia.

Haeberli, Eric. 2002. Observations on the Loss of V2 in English. In *Studies in Comparative Syntax*, eds. C. Jan-Wouter Zwart and Werner Abraham, 245–272. Amsterdam: John Benjamins.

Heycock, Caroline and Antonella Sorace. To appear. Verb movement in Faroese: New perspectives on an old problem. In *Nordlyd 35: Solf Proceedings*, ed. by Øystein Alexander Vangsnes, Gunnar Hrafn Hrafnbjargarson, and Christine Bjerkan Østbø. *Tromsø Working Papers in Language and Linguistics*.

Holmberg, Anders and Christer Platzack. 1995. *The Role of Inflection in Scandinavian Syntax*. New York: Oxford University Press.

Hrafnbjargarson, Gunnar Hrafn. 2004. *Oblique Subjects and Stylistic Fronting in the History of Scandinavian and English: The Role of IP-Spec*. Ph.D. thesis, University of Århus. Available at http://folk.uio.no/gunnahh/phd/.

Hróarsdóttir, Thorbjörg, Kristine Bentzen, Gunnar Hrafn Hrafnbjargarson, and Anna-Lena Wiklund. 2007. The afterglow of verb movement. *Working Papers in Scandinavian Syntax* 80:45–75.

Jonas, Dianne. 2002. Residual V-to-I. In *Syntactic Effects of Morphological Change*, ed. David W. Lightfoot, 251–270. Oxford: Oxford University Press.

Labelle, Marie. 2007. Clausal architecture in Early Old French. *Lingua* 117:289–316.

Larsson, Carl. 1931. *Ordföljdsstudier över det finita verbet i de nordiska fornspråken [Word order studies of the finite verb in Old Scandinavian]*. Uppsala: Lundequistska.

Ledgeway, Adam. 2008. Satisfying V2 in early Romance: Merge vs. Move. *Journal of Linguistics* 44:437–470.

Lightfoot, David. 1999. *The Development of Language: Acquisition, Change, and Evolution*. Malden, Ma and Oxford: Blackwell.

Mikkelsen, Kristian. 1911. *Dansk ordföjningslære*. Copenhagen: Lehmann & Stage. Reprinted 1975.

Nygaard, Marius. 1906. *Norrøn Syntax*. Kristiania: H. Aschehoug.

Petersen, Hjalmar. 2000. IP or TP in Modern Faroese. *Working Papers in Scandinavian Syntax* 66:75–83.

Pintzuk, Susan. 1999. *Phrase Structure in Competition: Variation and Change in Old English Word Order*. New York/London: Garland. Reprint with minor revisions of 1991 University of Pennsylvania PhD dissertation.

Platzack, Christer, and Anders Holmberg. 1989. The role of AGR and finiteness. *Working Papers in Scandinavian Syntax* 43:51–76.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of Grammar: Handbook in Generative Syntax*, ed. Liliane Haegeman, 281–337. Dordrecht: Kluwer.

Roberts, Ian. 1993. *Verbs and Diachronic Syntax: A Comparative History of English and French*. Dordrecht: Kluwer.

Roberts, Ian. 1996. Remarks on the Old English C-system and the diachrony of V2. *Linguistische Berichte* 7:154–168. Sonderheft.

Rohrbacher, Bernhard. 1999. *Morphology-Driven Syntax: A Theory of V to I Raising and Pro-Drop*. Amsterdam: John Benjamins.

Santorini, Beatrice and Anthony Kroch. 2007. The syntax of natural language. Manuscript, University of Pennsylvania.

Sundquist, John D. 2002. Object Shift and Holmberg's Generalization in the history of Norwegian. In *Syntactic Effects of Morphological Change*, ed. David W. Lightfoot, 326–347. Oxford: Oxford University Press.

Sundquist, John D. 2003. The Rich Agreement Hypothesis and Early Modern Danish embedded-clause word order. *Nordic Journal of Linguistics* 26 2: 233–258.

Thráinsson, Höskuldur, Hjalmar P. Petersen, Jógvan í Lon Jacobsen, and Zakaris Svabo Hansen. 2004. *Faroese: An Overview and Reference Grammar*. Føroyaðskaparfelag.

Thráinsson, Höskuldur. 2003. Minimalist syntax. In *Syntactic Variation, Historical Development and Minimalism*, ed. Randall Hendrick, 152–191. Oxford: Blackwell.

van Kemenade, Ans. 1987. *Syntactic Case and Morphological Case in the History of English*. Foris: Dordrecht.

Vance, Barbara. 1989. *Null Subjects and Syntactic Change in Medieval French*. Ph.D. thesis, Cornell University, Ithaca, N.Y.

Vance, Barbara. 1997. *Syntactic Change in Medieval French: Verb-Second and Null Subjects*. Dordrecht: Kluwer.

Vikner, Sten. 1995. *Verb Movement and Expletive Subjects in the Germanic Languages*. New York: Oxford University Press.

Warner, Anthony. 2007. Parameters of variation between verb-subject and subject-verb order in late Middle English. *English Language and Linguistics* 11 1: 81–111.

Westergaard, Marit and Øystein Alexander Vangsnes. 2005. Wh-questions, V2, and the left periphery in three Norwegian dialects. *Journal of Comparative Germanic Linguistics* 8:117–158.

Westergaard, Marit. 2005. *The Development of Word Order in Norwegian Child* Language. Ph.D. thesis, University of Tromsø, Norway.

Westergaard, Marit. 2009. Word order in Old and Middle English: The role of information structure and first language acquisition. *Diachronica* 26 1:65–102.

Wiklund, Anna-Lena, Gunnar Hrafn Hrafnbjargarson, Kristine Bentzen, and Thorbjörg Hróarsdóttir. 2007. Rethinking Scandinavian verb movement. *Journal of Comparative Germanic Linguistics* 10 3:203–233.

Zanuttini, Raffaella. 1997. Negation and Clausal Structure. A Comparative Study of Romance Languages. New York: Oxford University Press.

Demystifying the Predicate Cleft Construction[*]

Jason Kandybowicz
Swarthmore College, Pennsylvania

Abstract

Certain varieties of predicate cleft constructions appear to derive from movement operations unobserved elsewhere in natural language. The dependencies that characterize these varieties are *wh*-like in terms of the locality constraints that limit the distances they may span, yet are unlike *wh*- constructions in two respects: they fail to leave gaps and do not target maximal projections. On the force of evidence from Nupe, I argue that predicate cleft constructions of this heretofore mysterious variety involve the formation of parallel chains, dependencies with overlapping tails and disjoint heads. I show that the construction's infamously recalcitrant properties can be derived as a consequence and thus, that the movements involved are neither exotic nor theoretically problematic.

1. Introduction

Loosely characterized, the predicate cleft construction (PCC) is a configuration in which a predicate is promoted in discourse prominence and doubly realized in peripheral and clause-internal positions. Crosslinguistically, the construction is attested in a wide range of unrelated language familes, though most associate it with the languages and creoles of West Africa and the Atlantic. Research into the construction has progressed for well over thirty years (see, for example, Awóbùluyi 1971 and Bamgbose 1972 for early accounts), leading to a number of new and important insights into movement theory, clausal architecture, and information structure, to name just a few areas. However, despite the considerable body of literature this research has produced and the careful attention PCCs have received on an analytical level, an air of mystery still surrounds the construction. It is this mystique that both obscures its analysis and gives it theoretical allure. The PCC's mystique owes largely to the nature of the

dependency between the predicates that characterizes the construction in a number of languages. On the assumption that this dependency is movement-based, as has become standard in the literature (although see Lumsden & Lefebvre 1990, Lumsden 1990, Larson & Lefebvre 1991, and Dekydspotter 1992 for alternative views), the movements involved in PCC formation in certain languages appear to be unlike those found elsewhere in natural language. For example, PCC dependencies cannot span distances of arbitrary lengths. In fact, they are constrained by the same principles of locality that govern *wh*-movement, as we'll see. At the same time, however, the movements are unlike the *wh-* variety in two respects. One, they systematically appear not to leave gaps (i.e. two occurrences of the same predicate are pronounced). And two, in some languages the movement operation does not seem to target maximal projections (*i.e.* the moved elements appear to be heads).

Of these gross characteristics, the latter property distinguishes predicate cleft constructions of the "better-understood" variety from those of the still "mysterious" variety. In languages that tolerate pied piping of the cleft (*e.g.,* Yoruba, Buli, Hebrew, and Russian, to name just a few), the movements are less mysterious than those in languages that do not allow predicate pied piping for the simple reason that the movements involved must be phrasal (or phrasal remnants). Examples of this "better-understood" class of predicate cleft construction appear below.

(1)   a. YORUBA (Kandybowicz 2004)
          Rí-ra   ni Olú ra ife.
          RED-buy FOC Olu buy cup
         'Its BUYING that Olu did to the cup.'

        b. **Rí-ra   ife** ni Olú ra ife.
          RED-buy cup FOC Olu buy cup
          'Its BUYING A CUP that Olu did.'

        c. BULI (Hiraiwa 2005a)
          Dē̄-kā   àl̄ī/ àtì Àtìm dɛ̀ mángò-kǔ dīēm.
          eat-NOML COMP Atim   ate mango-D yesterday
          'It is eating that Atim ate the mango yesterday.'

        d. **Mángò-kú dē̄-kā**   àl̄ī/ àtì Aòtìm dɛ̀ dīēm.
          mango-D   eat-NOML COMP Atim   ate yesterday
          'It is eating the mango that Atim ate yesterday.'

    e. HEBREW (Landau 2006)

      Le'ha'amin, hu ma'amin be-nisim.

      believe$_{INF}$ he believes in-miracles

      'As for believing, he believes in miracles.'

    f. **Le'ha'amin be-nisim**,   hu ma'amin.

      believe$_{INF}$ in-miracles he believes

      'As for believing in miracles, he believes.'

    g. RUSSIAN (Abels 2001)

      Dumat´ o ž   ž enit´be -to on dumaet, no nikogda on ne ž enitsja.

      think$_{INF}$ about marriage TO he thinks but never he not marry-self

      'He does think about marriage, but he will never marry.'

The only salient property differentiating the movements in pied piped PCCs from those of the standard *wh*- variety, then, is the existence of missing gaps.[1] This property, however, has recently been explored to a considerable extent within the framework of the Copy theory of movement (*e.g.,* Abels 2001, Hiraiwa 2005a, Landau 2006, among others), which is why PCCs of this variety can be referred to as "better-understood". Predicate clefts of the other variety, however, continue to be regarded as mysterious because in addition to the existence of missing gaps, heads appear to move like *wh*- phrases. One language of this "mysterious" variety is Nupe, a Benue-Congo language spoken in Nigeria. The preliminary data presented below illustrate the basic properties of PCC movement in the language. (More detailed paradigms follow in later sections.) (2a) shows that unlike *wh*- movement, PCC movement in Nupe fails to leave a gap. (2b) shows that the cleft element is non-phrasal. And (2c) illustrates that PCC movement in the language is island sensitive, just as *wh*- movement is. Note furthermore that although lexically related, the pronounced occurrences of the predicate differ morphologically, as is typical in predicate clefting languages (cf. (1)).

(2)   a. **Si-si**   Musa à *(**si**) nakàn o.

        RED-buy Musa FUT buy meat FOC

        'It is BUYING that Musa will do to the meat (as opposed to say, *selling*).'

     b. *****Si-si**   **nakàn** Musa à   **si** (**nakàn**) o.

        RED-buy meat   Musa FUT buy meat   FOC

        'It is BUYING MEAT that Musa will do.'

---

[1] This is not meant to deny the existence of missing gaps in *wh*- constructions. Partial *wh*-movement constructions, in which a *wh*- constituent is doubly realized, have been attested in a variety of languages. What differentiates the existence of missing gaps in *wh*- movement constructions from those in PCCs is that in the former, missing gaps are the exception, whereas in the latter, they are the rule.

c. *Si-si Musa gbíngàn [ké Gana **si** o] o.
 RED-buy Musa ask what Gana buy FOC FOC
 'Musa asked what Gana BOUGHT.'

As such, the movements that characterize PCCs of the "mysterious" variety seem to be of the "A-bar/*wh*- head movement" kind, as characterized by Koopman (1984). In this article, based on an investigation of Nupe PCCs, I argue that there is nothing mysterious or exotic about PCC movements of this variety and that properties that seemed mystifying in the past (i.e. *wh*- head movement/missing gaps) can be derived and demystified once certain Minimalist assumptions are adopted, namely, that chains may be formed in parallel and heads and phrases pattern alike with respect to movement.

2. Core Properties of Nupe Predicate Cleft Constructions

2.1 Predicate vs. Non-predicate Focus

It will be instructive to first consider how predicate clefting differs from other instances of focus in the language. In addition to predicates, a variety of constituents may be focused in Nupe (3b-d). In all cases, the focused element appears in a left-peripheral position and the focus marker surfaces clause-finally. Unlike predicate focus (3e), though, the focused element's morphological form does not change and the movement clearly leaves a gap in its extraction site. These facts are presented below.

(3) a. NEUTRAL SENTENCE
   Musa à ba nakàn sasi èsun làzì yin[2].
   Musa FUT cut meat some tomorrow morning PRT
   'Musa will cut some meat tomorrow morning.'

  b. SUBJECT FOCUS

   Musa ___ à ba nakàn sasi èsun làzì yin o.
   Musa FUT cut meat some tomorrow morning PRT FOC
   'MUSA will cut some meat tomorrow morning.'

  c. OBJECT FOCUS

   Nakàn sasi Musa à ba ___ èsun làzì yin o.
   meat some Musa FUT cut tomorrow morning PRT FOC
   'Musa will cut SOME MEAT tomorrow morning.'

---

[2] The *yin* particle that surfaces in these examples is a temporal adverbial particle.

d. MODIFIER FOCUS

   Èsun   làzì   Musa à   ba nakàn sasi ___ yin o.
   tomorrow morning Musa FUT cut meat some   PRT FOC
   'Musa will cut some meat TOMORROW MORNING.'

e. PREDICATE FOCUS

   Bi-ba   Musa à *(ba) nakàn sasi èsun   làzì   yin o.
   RED-cut Musa FUT cut meat some tomorrow morning PRT FOC
   'It is CUTTING that Musa will do to some meat tomorrow morning.'

Non-predicate focus thus appears to involve A-bar chain formation (more on this in the following sub-section), formed by extraction of the focused constituent and the PF deletion of its tail, as in typical instances of chain formation. This analysis, however, does not appear to straightforwardly extend to PCC formation in the language.

2.2 Duality of Movement

As discussed in the introduction, the theoretical allure of non pied piping PCCs is that they appear to involve movement operations whose properties are otherwise unattested elsewhere in natural language. Piou (1982) and Koopman (1984) first observed that PCCs in unrelated languages (Haitian and Vata, respectively) are *wh*-like in that the distances they may span are constrained, yet at the same time, they are unlike *wh-* constructions in that they appear not to leave a gap or target a maximal projection. Thus, PCCs of this variety seem to necessitate the admission of a third displacement type into the movement typology, i.e., one that is intermediate between head movement and phrasal movement. The challenge posed by this variety of PCC, then, is to explain why it behaves like A-bar movement in some respects, but not in others.

   Nupe PCCs seem to warrant the same conclusions that Piou and Koopman drew. Although the dependency between the focused left peripheral predicate and the lower occurrence is unbounded, crossing finite clause boundaries in the presence of bridge verbs (4a), it is also island sensitive (4b-i), both hallmarks of A-bar dependencies.

(4)   a. √ EXTRACTION ACROSS THE CLAUSAL COMPLEMENT OF A BRIDGE VERB
          Si-si   Musa gàn [gànán Nànǎ kpe [gànán Gana **si**   eci]] o.
          RED-buy Musa say   COMP Nana know COMP Gana buy yam FOC
          'It was BUYING that Musa said that Nana knows that Gana did to a yam.'

b. *EXTRACTION ACROSS THE CLAUSAL COMPLEMENT OF A NON-BRIDGE VERB

   *Si-si   u: tán Musa [gànán mi: **si**   doko] o.
   RED-buy 3RD.SG pain Musa COMP 1ST.SG buy horse FOC
   'It pained Musa that I BOUGHT a horse.'

c. *WH*-ISLAND

   *Si-si   Musa gbíngàn [ké   Gana **si**   o]   o.
   RED-buy Musa ask   what Gana buy FOC FOC
   'Musa asked what Gana BOUGHT.'

d. COMPLEX NP ISLAND

   *Gi-gi   Musa si [bise na   **gi** eyì   na] o.
   RED-eat Musa buy hen   COMP eat corn PRT FOC
   'Musa bought the hen that ATE the corn.'

e. SUBJECT ISLAND

   *Si-si   [gànán etsu **si**   doko] tán Musa o.
   RED-buy COMP chief buy horse pain Musa FOC
   'That the chief BOUGHT a horse pained Musa.'

f. ADJUNCT ISLAND[3]

   *Bi-ba   [Musa gá   è   **ba** nakàn] o,   Gana à   pa   eci.
   RED-cut Musa COND PRS cut meat FOC Gana FUT pound yam
   'If Musa is CUTTING the meat, then Gana will pound a yam.'

g. Musa gá è ba nakàn, **pi-pa** [Gana à   **pa**   eci] o.
   Musa COND PRS cut meat, RED-pound Gana FUT pound yam FOC
   'If Musa is cutting the meat, then it is POUNDING that Gana will do to a yam.'

h. COORDINATE ISLAND[4]

   *Bi-ba   [Musa$_i$ à   **ba** nakàn] u:$_i$ ma à   du   cènkafa o.
   RED-cut Musa FUT cut meat   3RD.SG and FUT cook rice   FOC
   'It is CUTTING that Musa$_i$ will do to the meat and he$_i$ will cook the rice.'

i. ***Du-du**   Musa$_i$ à ba nakàn [u:$_i$ ma à   **du** cènkafa] o.
   RED-cook Musa FUT cut meat 3RD.SG and FUT cook rice   FOC
   'Musa$_i$ will cut the meat and it is COOKING that he$_i$ will do to the rice.'

---

[3] As in English, the CP containing the conditional marker *gá* (i.e. the antecedent) is an adjunct, while the consequent CP is the host. This is confirmed by the PCC extraction asymmetry in (4f-g).

[4] ATB movement of both predicates is ungrammatical in coordinate structures as well:

(i)   *Bi-ba   du-du   Musa$_i$ à ba nakàn u:$_i$ ma à   du   cènkafa o.
         RED-cut RED-cook Musa FUT cut meat   3RD.SG and FUT cook rice   FOC
          'Musa$_i$ will CUT the meat and he$_i$ will COOK the rice.'

In addition, PCCs and *wh-* questions are in complementary distribution, as shown below. This further suggests the A-bar status of both constructions – focused constituents and *wh-* DPs compete for the same left peripheral focus position.

(5)   a.   *Ké   bi-ba   Musa ba o?
              what RED-cut Musa cut FOC
              'What did Musa CUT?'

        b.   *Bi-ba   ké   Musa ba o?
              RED-cut what Musa cut FOC

Despite their affinity to *wh-* constructions, Nupe PCCs exhibit properties that distinguish themselves from *wh-*/phrasal movement constructions, as alluded to in the introduction. Let's concentrate on how the two constructions differ.

As previously mentioned, if PCCs involve predicate extraction, they are unlike typical instances of *wh-* movement in that multiple links of their chains are consistently realized. That is, whereas standard *wh-* movement leaves a gap, predicate focus leaves behind a "resumptive verb" in Koopman's (1984) words. In addition, although *wh-* elements can be focused *in situ* in many languages, predicate focus in Nupe can only be achieved when the predicate appears in the left periphery.

(6)   a.   *Musa **bi-ba**   ba nakàn o.
              Musa RED-cut cut meat   FOC

        b.   *Musa ba **bi-ba**   nakàn o.
              Musa cut RED-cut meat   FOC

        c.   *Musa ba nakàn **bi-ba**   o.
              Musa cut meat   RED-cut FOC

        d.   *Musa ba nakàn o   **bi-ba**.
              Musa cut meat   FOC RED-cut

Another important difference between Nupe PCCs and *wh-*/focus constructions is the fact that the latter clearly involve left-peripheral *phrasal* constituents (cf. (3c-d)), while in the former, the peripheral elements appear to be heads. Similar to languages like Vata and Haitian, but unlike Yoruba, Buli, Russian, and Hebrew, the verb's arguments cannot appear in the left periphery with the focused predicate (7a,b). In fact, Nupe is more conservative than Vata and Haitian because tense markers, aspectual elements, and low adverbs, which can accompany the cleft element in both languages, are restricted from appearing at the left edge of the clause with the focused predicate (7c-e). And unlike a number of genetically-related Nigerian languages (for instance, Edo), the focused

predicate may not appear with nominal modifiers (7f-g). Thus, unlike *wh*-movement, the cleft predicate in Nupe is unable to pied pipe material.

(7)   a. *[Du-du   cènkafa] Musa à   du (cènkafa) o.
RED-cook rice   Musa FUT cook rice   FOC
'It is COOKING RICE that Musa will do.'

b. *[Cènkafa du-du]   Musa à   du (cènkafa) o.
rice   RED-cook Musa FUT cook rice   FOC
'It is COOKING RICE that Musa will do.'

c. *[(à) du-du   (à)] Musa à   du   cènkafa o.
FUT RED-cook FUT Musa FUT cook rice   FOC
'It is COOKING that Musa will do to the rice.'

d. *[(á) du-du   (á)] Musa á   cènkafa du   o.
PRF RED-cook PRF Musa PRF rice   cook FOC
'It is COOKING that Musa has done to the rice.'

e. *[(Dàdà) du-du   (sanyín)] Musa à   du   cènkafa o.
quickly RED-cook quietly Musa FUT cook rice   FOC
'It is QUICK/QUIET COOKING that Musa will do to the rice.'

f. *[Wu-wu gútá] Gana wu Musa o.
RED-hit three Gana hit Musa FOC
'It was HITTING THREE TIMES that Gana did to Musa.'

g. *[Wu-wu wangi] Gana wu Musa o.
RED-hit good Gana hit Musa FOC
'It was A GOOD HITTING that Gana gave to Musa.'

This restriction on pied-piping is limited to PCCs. Pied-piping is independently attested in the language, as the data below indicate.

(8)   a.   Ké Musa ba nakàn bè ___ yin o?
what Musa cut meat with   PRT FOC
'What did Musa cut the meat with?'

b.   Bè   ké Musa ba nakàn ___ yin o?
with what Musa cut meat   PRT FOC
'With what did Musa cut the meat?'

It is tempting, therefore, to analyze the Nupe cleft predicate as a kind of deverbal head. Ultimately, this is the source of the duality of predicate focus movement in the language. With respect to locality and complementarity with *wh-* questions, Nupe predicate focus patterns like A-bar movement. Yet at the

same time, the predicate focus chains appear neither to be reduced nor obviously headed by a maximal projection. Furthermore, if the dependency between the focused predicate and the matrix verb arises as a consequence of chain formation, it is not immediately apparent why there is a morphological difference between the two elements. The adequacy of any Nupe PCC analysis can thus be judged by how well it resolves these descriptive and theoretical tensions.

3. Analysis and Derivation

3.1 Extraction of Heads, not Phrases

The view that PCC derivations involve phrasal fronting is common to many analyses (Nishiyama & Cho 1998, Koopman 1999, Cho & Nishiyama 2000, Abels 2001, Nunes 2003, 2004, Hiraiwa 2005a, Landau 2006, among others). In many languages that allow predicate clefting, the displaced predicate may pied pipe verbal arguments (*cf.* Yoruba (Cho & Nishiyama 2000), Buli (Hiraiwa 2005a), Russian (Abels 2001), Hebrew (Landau 2006)), tense/aspect markers (*cf.* Vata (Koopman 1984), Haitian (Piou 1982)), certain types of modifiers (*cf.* Edo (Stewart 2001), Vata (Koopman 1984), Haitian (Piou 1982)), and even serialized verbal occurrences (*cf.* Yoruba (Baker 1989, Gruber & Collins 1996, Manfredi 1993, Cho & Nishiyama 2000), Dagaare (Bodomo 2004), Buli (Hiraiwa 2005a)). We have seen that the first three of these possibilities do not obtain in Nupe (*cf.* (7)).

The other cross-linguistically attested extraction pattern, namely, the pied piping of serialized predicates, is also unavailable. The data below illustrate that although V1 may undergo predicate cleft, V2 pied piping is impossible across all serial verb construction types in the language. (The names of the Nupe serial verb construction varieties shown below are taken from Stewart's (2001) typology.)

(9)   RESULTATIVE SERIAL VERB CONSTRUCTION CLEFT PATTERNS

    a. Fu-fo   Musa fo   èwò   li   o.
        RED-wash Musa   wash garment be clean FOC
        'It was WASHING that Musa washed the garment clean.'

    b. *[Fu-fo   **li-li**]   Musa fo   èwò   (li)   o.
        RED-wash RED-be clean Musa wash garment be clean FOC
        'It was WASHING CLEAN that Musa did to the garment.'

Jason Kandybowicz

CONSEQUENTIAL SERIAL VERB CONSTRUCTION CLEFT PATTERNS

c. Pi-pa   Musa pa   eci gí o.
   RED-pound Musa pound yam eat FOC
   'It was POUNDING and then eating that Musa did to a yam.'

d. *[Pi-pa   **gi-gí**]   Musa pa   eci (gí) o.
    RED-pound RED-eat Musa pound yam eat   FOC
    'It was POUNDING and then EATING that Musa did to a yam.'

PURPOSIVE SERIAL VERB CONSTRUCTION CLEFT PATTERNS

e. Si-si   Musa si eyì dzò o.
   RED-buy Musa buy corn plant FOC
   'It was BUYING that Musa did to the corn in order to plant it.'

f. *[Si-si   **dzu-dzò**]   Musa si   eyì (dzò) o.
    RED-buy eat RED-plant Musa buy corn plant FOC
    'It was BUYING to PLANT that Musa did to the corn.'

The generalization is clear. In Nupe PCCs, only the focused predicate is realized in the left periphery. Phrasal constituents are excluded. The logical conclusion is thus that Nupe PCC derivations involve the extraction of heads, not phrases. Of course, it is still analytically possible to maintain that the focused constituent is a remnant phrasal category containing both the predicate and copies of the scrambled verbal arguments (cf. Koopman 1999, Abels 2001, Nunes 2003). The burden of proof, however, would be to explain why these arguments are systematically unable to be phonetically realized alongside the focused predicate, a possibility that arises once the Copy theory of movement is assumed. (See Vicente 2009 on this point.) Given these considerations, a phrasal movement analysis does not seem motivated in the case of Nupe PCCs. We are left with the conclusion that PCC formation in the language is driven by head movement.

3.2 The Derivation of Nupe Predicate Cleft Constructions

I have argued that the movement operation responsible for predicate focus in Nupe is of the head movement variety. This entails the movement of an X^{MIN} category into the specifier position of Focus Phrase (cf. Koopman's (1984) conception of "A-bar head movement" and Collins & Essizewa's (2007) analysis of verb phrase-internal focus in Kabiye). This conclusion, however, flies in the face of a standard assumption in Generative Grammar, namely, that movement operations affect heads and phrases differently. Ultimately, this idea is a relic of Emonds' (1970) Structure Preservation hypothesis, which was carried over into X-bar theory. With the elimination of the X-bar theoretic conception of structure

building in Chomsky's (1995) Bare Phrase Structure theory, the distinction between head and phrasal movement was revived by means of Chomsky & Lasnik's (1993) Chain Uniformity condition, which prohibits chain links from having differing X^{min}/X^{max} statuses. To the extent that the Chain Uniformity condition is real, the analysis of Nupe PCCs I am proposing is conceptually problematic. Vicente (2005, 2006), however, argues convincingly that the Chain Uniformity condition is untenable on the grounds that a) it is conceptually suspect under Minimalist desiderata, b) it has a limited/inconsistent domain of application, and c) its effects can be independently derived. If true, the elimination of both X-bar theory and the Chain Uniformity condition removes any conceptual barrier barring head movement into specifier positions. My head movement analysis of Nupe PCCs is guided by these considerations.

In this section, I divide my analysis of Nupe PCCs into two separate issues: the bi-locational realization of the predicate root and the allomorphy of the phonetically realized predicate occurrences.

3.2.1 Deriving the Bi-locational Realization of the Predicate

Independent of PCC formation, Nupe predicate roots (represented below as √) raise to v^0 (Kandybowicz & Baker 2003, Kandybowicz 2008a, 2009), where they are assigned verbal features by the verbalizing morpheme (v^0), as is the case in many languages.

The focused predicate, however, surfaces with nominal features (instantiated by its reduplicative morphology, as is common in the languages of West Africa). The data below support the claim that reduplicated verbs are categorically nominal. Observe that reduplicated predicates may appear in both subject and object gerunds.

(10)

vP
├── v
│   ├── √ⁱ
│   └── v
└── √P
    └── …√ⁱ…
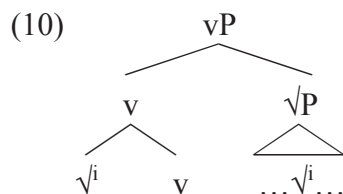
(11) a. [Bi-ba   na   u: ba nakàn na] tan Musa.
          RED-cut COMP 3rd.SG cut meat   PRT pain Musa
          'His cutting the meat pained Musa.'

      b. Musa sundàn [bi-ba   nyá Gana].
          Musa fear   RED-cut POSS Gana
          'Musa feared Gana's cutting.'

Jason Kandybowicz

Given that the focused predicate is nominal, it follows that the cleft predicate root excludes the verbalizing morpheme (v^0). This means that the disp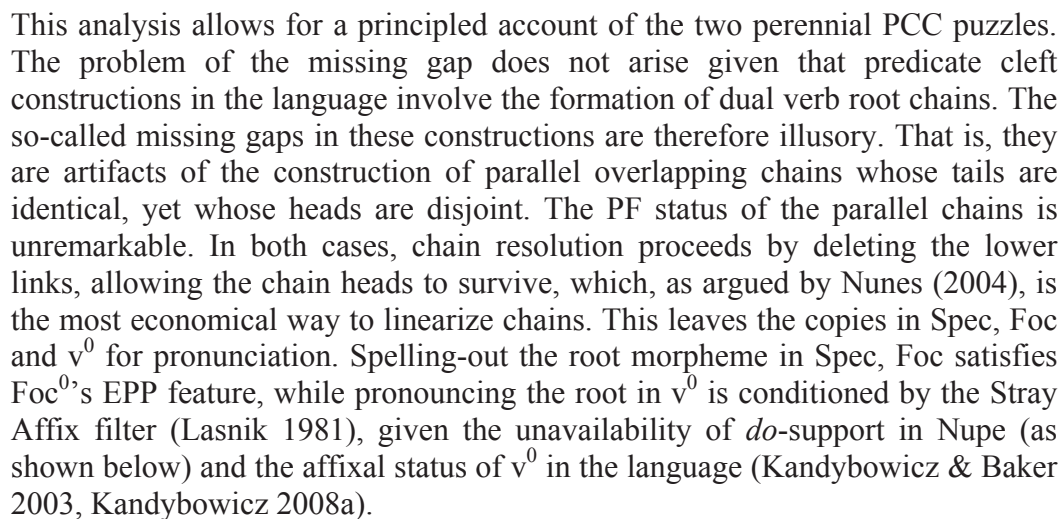laced left peripheral root cannot be a link in a chain that also includes the positions related by $\sqrt{}$-to-v^0 head movement. In other words, the predicate root does not cyclically raise to Spec, Foc after first raising to v^0. This is a welcome conclusion considering that such an instance of chain formation would involve sub-extraction out of a complex head (*i.e.,* excorporation). I illustrate this graphically below.

(12)
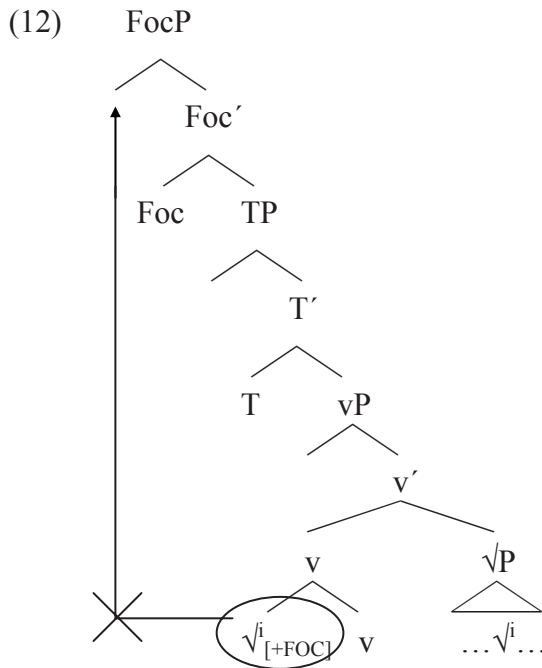
Consequently, Nupe PCC derivations must involve the formation of two distinct (*i.e.,* parallel) root chains; one chain formed by head raising $\sqrt{}$ to v^0 (as in simple declaratives) and a separate chain formed by cyclically raising $\sqrt{}$ to Spec, Foc via the edge of the vP phase (see (13) below). Aboh (2006) and Aboh & Dyakonova (2009) propose analyses in this spirit for predicate cleft constructions in Gungbe and Russian respectively. The formation of these two chains, I claim, is a consequence of the fact that the focus feature-bearing $\sqrt{}$ morpheme is probed by both v^0 and Foc^0. My proposal is laid out schematically below.

(13)

This analysis allows for a principled account of the two perennial PCC puzzles. The problem of the missing gap does not arise given that predicate cleft constructions in the language involve the formation of dual verb root chains. The so-called missing gaps in these constructions are therefore illusory. That is, they are artifacts of the construction of parallel overlapping chains whose tails are identical, yet whose heads are disjoint. The PF status of the parallel chains is unremarkable. In both cases, chain resolution proceeds by deleting the lower links, allowing the chain heads to survive, which, as argued by Nunes (2004), is the most economical way to linearize chains. This leaves the copies in Spec, Foc and v^0 for pronunciation. Spelling-out the root morpheme in Spec, Foc satisfies Foc^0's EPP feature, while pronouncing the root in v^0 is conditioned by the Stray Affix filter (Lasnik 1981), given the unavailability of *do*-support in Nupe (as shown below) and the affixal status of v^0 in the language (Kandybowicz & Baker 2003, Kandybowicz 2008a).

(14)   a. Bi-ba   Musa ba nakàn o.
             RED-cut Musa cut meat   FOC
             'It was CUTTING that Musa did to the meat.'

        b. *Bi-ba   Musa **dzin** nakàn o.
             RED-cut Musa do   meat   FOC

Jason Kandybowicz

Because exactly two root chains are formed and no ancillary morphological or phonological requirements exist to force the spell-out of a lower chain link in this construction, the number of phonetically realized predicates in a Nupe PCC is limited to two occurrences. The bi-locational realization of the root morpheme and thus the appearance of missing gaps follows.

The issue of the duality of movement derives from the fact that a head is moving like a phrase with respect to the target of movement. In short, the movement *does* in fact have a dual nature. However, if one is persuaded by Vicente's (2005, 2006) elimination of the Chain Uniformity condition, the "problem" of *wh-* head movement can ultimately be regarded as a non-issue. Thus, we predict to find similar movement patterns in the languages of the world. See Vicente 2006, Collins & Essizewa 2007 and Harbour 2008 for a variety of examples, lending support to the existence of this variety of movement.
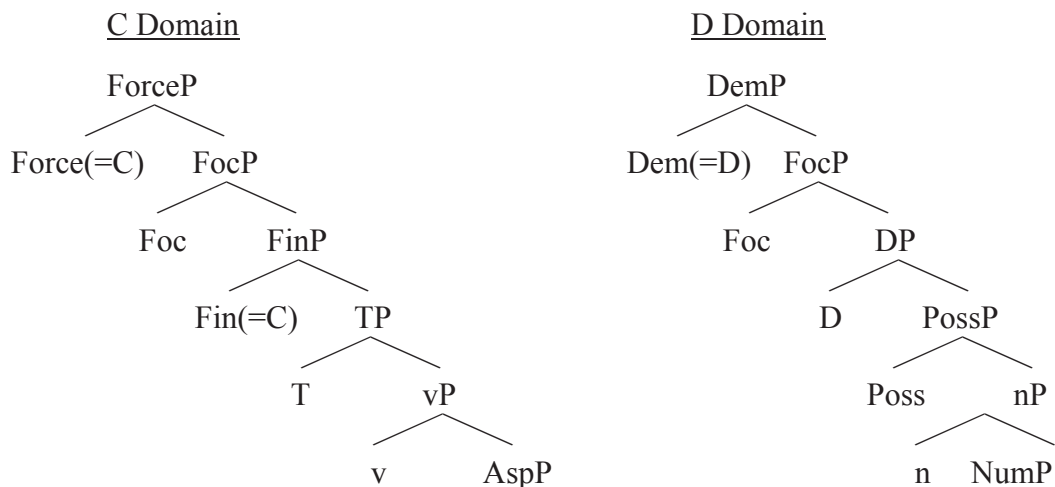
3.2.2 Deriving the Predicate Root Allomorphy

In Distributed Morphology (DM), the morphological construction of words is constrained by the hierarchical structures assembled in the narrow syntax. Thus, words do not enter derivations pre-formed. Under the DM approach, abstract morphemes (terminals) manipulated in the syntax are underspecified for various grammatical properties. Root morphemes, in particular, are underspecified for syntactic category. These category-neutral pieces are assigned categorial features in virtue of occupying positions that lie under the scope of category-assigning functional morphemes at the point of Vocabulary Insertion. In DM, Vocabulary Insertion is a derivationally late operation occurring after the syntactic computation that provides terminal nodes with phonetic content. To concretize this description a bit, a root morpheme under the scope of a head bearing verbal features (*e.g.,* v^0) will surface with the category feature [+V] and will subsequently be spelled out as a verb. Under the scope of a head with nominal features (*e.g.,* D^0/n^0), on the other hand, the same root will inherit the feature [+N] and will surface with nominal morphology (*i.e.,* as a nominalization). This conception of the syntax-morphology interface allows for an elegant and theoretically principled analysis of the morphological mismatch between the peripheral (focused) predicate and the lower verbal occurrence in Nupe PCCs.

Given that v^0 is not pied-piped with the focused $\sqrt{}$ morpheme, as argued earlier, the displaced term in a Nupe PCC is not one whose category status is established. Rather, what has moved is an element that awaits categorial determination. The nominal features borne by the focused predicate, I claim, are environmentally conditioned. I appeal here to the discovery that the C domain (i.e. the left periphery) is a nominal domain on a parallel with the D domain

(Chomsky 1970, Abney 1987, Lefebvre & Muysken 1988, Szabolcsi 1994, Collins 2001, Hiraiwa 2005a,b, among others).

(15) CP/DP PARALLELISM (Hiraiwa 2005a,b)

<div align="center">

C Domain   D Domain

ForceP   DemP

Force(=C) FocP   Dem(=D) FocP

Foc FinP   Foc DP

Fin(=C) TP   D PossP

T vP   Poss nP

v AspP   n NumP

</div>

As such, in virtue of its movement to a nominal domain (i.e. a position under the scope of C^0), the predicate root is assigned nominal features and spelled out in a reduplicated (nominalized) form following Vocabulary Insertion. Because the lower root copy is adjoined to v^0, it is spelled out with verbal morphology. The morphological mismatch between the two pronounced predicates thus follows. Space restrictions preclude a more detailed discussion of the nominalization of the focused root. See Kandybowicz 2008b for a more detailed analysis.

4. Parallel Chain Formation

I have appealed to the concept of parallel chain formation as a way of demystifying the movement operation responsible for predicate clefts of the Nupe variety. In this section, following a brief discussion of the concept itself, I review how the mysterious aspects of PCC movement in Nupe can be derived from the assumption that PCC derivations involve parallel chain formation.

      Chomsky (2008) motivates the existence of parallel chain formation without explicitly referring to it as such. For him, parallel chains emerge as a consequence of the existence of phase heads bearing two types of uninterpretable movement-triggering features, namely, Agree-features and Edge-features. If one of these features (the Agree-feature) is inherited by the lexical item selected by a phase head and both features probe for a common goal, then parallel chains will

Jason Kandybowicz

be formed. Chain$_1$ would relate the goal to the Edge-feature of the phase head, while chain$_2$ would relate the goal to the inherited Agree-feature borne by the lexical item selected by the phase head. Among other reasons, Chomsky's parallel chains analysis is designed to eliminate string-vacuous movements such as the triple linked successive cyclic A´-A-A chains typically assumed in cases of subject *wh-* movement. On this approach, *wh-* copies in Spec, C and Spec, T are unrelated; they are links of separate chains. For reasons related to pronunciational economy, Chomsky assumes that only the copy in Spec, C is pronounced (i.e. only one of the two parallel chains is given phonetic content at PF). Apart from the PF resolution of the parallel chains, Aboh (2006) and Aboh & Dyakonova's (2009) parallel chains analyses of PCCs in Gungbe and Russian follow similar lines. For concreteness, I illustrate Chomsky's parallel chains analysis of English subject *wh-* derivations below.

(16)   who^i [C$_{EDGE}$ [who^i [T$_{AGREE}$ who^i [v [see John]]]]]
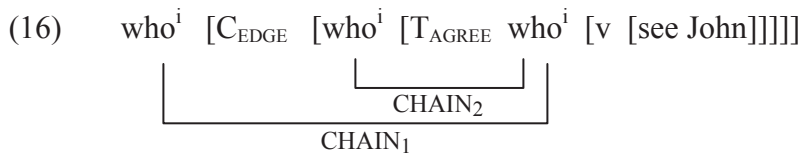
CHAIN$_2$

CHAIN$_1$

Chains 1 and 2 above are not necessarily formed at the same time. The chain formed by C^0's Edge-feature (chain$_1$) likely occurs first, given that the feature driving the formation of the T-*wh* chain is inherited after C^0 is merged. Thus, parallel chains are not necessarily chains that are created simultaneously. Rather, they are chains that are formed independently of one another.

The condition that triggers parallel chain formation is quite simple and does not necessarily require FEATURE INHERITANCE, as in Chomsky 2008. The minimal factor conditioning the formation of parallel chains is that separate Agree operations target one and the same goal. This is precisely the situation that obtains in Nupe PCC derivations. The v head probes for √, as does Foc^0. Parallel chain formation ensues.

Given these considerations, we can formalize the characterization of parallel chain formation in the following way. The statement below captures the fact that the formation of parallel chains involves distinct Agree operations with overlapping targets, as in Chomsky 2008, Aboh 2006, Aboh & Dyakonova 2009, and Collins & Essizewa 2007:198.

(17)   PARALLEL CHAIN FORMATION
        Two chains α and β are related by parallel chain formation if and only if:
        (i)   Tail (α) = Tail (β), AND
        (ii)   Head (α) ≠ Head (β)

As formalized above, conditions (i) and (ii) may be regarded as sufficient conditions for multiple copy pronunciation. Because copies of the same lexical item will be spread over independent/non-interacting chains, multiple pronunciation will result under ordinary circumstances of chain resolution. Again, this is precisely what we observe in the derivation of Nupe PCCs. The lower $\sqrt{}$ chain, internal to the v phase, is reduced in the normal way at PF – the chain head is spelled-out and the lower copy is deleted. The higher $\sqrt{}$ chain, part of the C phase, is linearized in the same way. Because the surviving $\sqrt{}$ copies occupy different phases, multiple copy realization does not pose a problem for linearization. In this way, we derive the appearance of missing gaps in Nupe PCCs. This is not the only "mysterious" property of Nupe PCC movement that can be derived under a parallel chains analysis. Recall that PCC movements in the language are puzzling in two other respects. One, heads pattern like phrases and two, the predicates that surface are obligatorily mismatched morphologically. The former is derivable under the unification of head and phrasal movement proposed by Vicente (2005, 2006), as previously discussed. The latter, however, also follows directly from the invocation of parallel chain formation, given the Late Insertion hypothesis of Distributed Morphology. Recall that parallel chains are defined as dependencies having overlapping tails and disjoint heads (17). In the case of parallel (category neutral) $\sqrt{}$ chains, this guarantees that the lexical content of the chains will overlap, but that the morphological/phonetic content of the chains will not. The latter property follows from two facts: the tail of any $\sqrt{}$ chain is categorically unspecified and the heads of parallel chains are disjoint (17ii). Because a) the categorical/morphological properties of $\sqrt{}$ chain links are determined by the time they occupy chain head positions and b) parallel chain heads are disjoint, the surviving $\sqrt{}$ occurrences will differ morphologically. In this way, we derive the allomorphy of predicates in PCCs from the assumption that PCC derivations of the Nupe variety involve the formation of parallel $\sqrt{}$ chains. This is the second respect in which parallel chain formation affords a way of demystifying the movement operation behind PCCs of the Nupe variety.

5. Summary and Closing Remarks

Predicate cleft constructions are by no means a homogeneous class. There are a number of different dimensions along which to taxonomize the construction, as others have pointed out (cf. Kandybowicz 2004). Because the focus of this article is chain formation, I have concerned myself primarily with the respects in which PCCs can be differentiated in terms of movement. I have divided PCCs into two broad classes accordingly: those in which the displaced predicate may pied pipe

Jason Kandybowicz

syntactic material and those in which the cleft is maximally a head. PCCs of the latter category (as exemplified in Nupe) exhibit what appear to be typologically exotic and theoretically problematic movement properties: the movement operation has the gross characteristics of *wh-* movement, yet unlike A-bar movement it systematically fails to leave gaps or target maximal projections. Because these properties seem to be observed only in PCCs, the movements that characterize PCCs of this variety have often been regarded as "mysterious". I have argued that the movements are neither exotic nor theoretically problematic on the force of the observation that the mysterious properties can be derived if heads are free to move like phrases and PCC derivations involve parallel chain formation. If correct, Nupe PCCs provide additional evidence for the inclusion of *wh-* head movement and parallel chain formation in Universal Grammar.

It is worth pointing out that PCCs of the Nupe variety are but one sub variety of the class of PCCs in which the cleft is maximally a head. In Nupe, the cleft predicate bears nominal morphology and nothing else. However, in other languages the cleft is more richly inflected. The parallel chains analysis developed in this article thus marries nicely to instances in which the cleft is morphologically impoverished (as in Nupe), but does not seem promising in the case of PCCs of the other sub variety, that is, for PCCs in which the cleft head bears (any amount of) inflectional morphology. In these cases, it is more plausible that PCC derivations involve the formation of a single head movement chain and that multiple chain links are phonetically realized at PF. The derivation of PCCs in terms of parallel vs. unary chain formation, if correct, would thus represent one way of characterizing the difference between predicate clefts of the Nupe sub variety and predicate clefts of the Vata/Haitian type within the broader classification of predicate cleft constructions formed by head movement.

References

Abels, Klaus. 2001. The predicate cleft construction in Russian. In *Formal approaches to Slavic Linguistics 9*, eds. Steven Franks, Tracy Holloway King and Michael Yadroff, 1-18. Ann Arbor: Michigan Slavic Publications.

Abney, Steven. 1987. The English noun phrase in its sentential aspect. Doctoral dissertation, MIT, Cambridge, MA.

Aboh, Enoch O. 2006. When verbal predicates go fronting. In *Papers on information structure in African languages*, eds. Ines Fiedler and Anne Schwarz, 21-48. Berlin: ZAS.

Aboh, Enoch O., and Marina Dyakonova. 2009. Predicate doubling and parallel chains. *Lingua* 119:1035-1065.

Awóbùluyi, Oladele. 1971. Splitting verbs in Yoruba. *Actes Du Hautieme De Linguistique Africanine* 1:151-164. Abidjan.

Baker, Mark C. 1989. Object sharing and projection in serial verb constructions. *Linguistic Inquiry* 20:513-553.

Bamgbose, Ayo. 1972. *The Yoruba verb phrase*. Ibadan: Ibadan University Press.

Bodomo, Adams. B. 2004. The syntax of nominalized complex verbal predicates in Dagaare. *Studia Linguistica* 58:1-22.

Cho, Eun, and Kunio Nishiyama. 2000. Yoruba predicate clefts from a comparative perpective. In *Advances in African Linguistics*. *Trends in African Linguistics 4*, eds. Vicki   Carstens and Frederick Parkinson, 37-49. Trenton: Africa World Press.

Chomsky, Noam. 1970. Remarks on nominalization. In *Readings in English transformational   grammar*, *eds.* Rederick Jacobs and Peter Rosenbaum, 184-221. Waltham, MA: Ginn & Company.

Chomsky, Noam. 1995. Bare phrase structure. In *Government and binding theory and the minimalist program*, ed. Gert Webelhuth, 383-439. Cambridge, MA: Blackwell.

Chomsky, Noam. 2008. On phases. In *Foundational issues in Linguistic theory*, eds. Robert Freidin, Carlos P. Otero and Maria Luisa Zubizarreta, 133-166. Cambridge, MA: MIT Press.

Chomsky, Noam and Howard Lasnik. 1993. The theory of principles and parameters. In *Syntax: An international handbook of contemporary research*, eds. Joachim Jacobs, Arnim   von Stechow, Wolfgang Sternefeld and Theo Vennemann, 506-569. Berlin: de Gruyter.

Collins, Chris. 2001. Aspects of plurality in úHoan. *Language* 77:456-476.

Collins, Chris, and Komlan Essizewa. 2007. The syntax of verb focus in Kabiye. In *Selected proceedings of the 37 annual conference on African Linguistics*, eds. Doris L. Payne   and Jaime Peña, 191-203. Somerville, MA: Cascadilla Press.

Dekydtspotter, Laurent. 1992. The syntax of predicate clefts. In *Proceedings of the twenty-second annual meeting of the north east linguistic society*, ed. Kimberely Broderick, 119-133. Amherst, MA: GLSA.

Donati, Caterina. 2006. On *wh-* head movement. In *Wh-movement: moving on*, eds. Lisa Cheng and Norbert Corver, 21-46. Cambridge, MA: MIT Press.

Emonds, Joseph. 1970. Root and structure preserving transformations. Doctoral dissertation, MIT, Cambridge, MA.

Gruber, Jeffrey, and Chris Collins. 1996. Argument projection, thematic configurationality and case theory. In *Projections and interface conditions: Essays on modularity*, ed. Anne-   Marie Di Sciullo, 130-154. Oxford: Oxford University Press.

Halle, Morris, and Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In   *The view from building 20*: *Essays in honor of Sylvain Bromberger*, eds. Kenneth Hale   and Samuel Jay Keyser, 111-176. Cambridge, MA: MIT Press.

Harbour, Daniel. 2008. Klivaj predika or predicate clefts in Haitian. *Lingua* 118:853-871.

Hiraiwa, Ken. 2005a. Predicate clefts in Buli: Categories and phases. *Linguistic Analysis* 32:544-583.

Hiraiwa, Ken. 2005b. Dimensions of symmetry in syntax: Agreement and clausal architecture. Doctoral dissertation, MIT, Cambridge, MA.

Kandybowicz, Jason. 2004. Predicate clefts, derivations, and universal grammar. In *Proceedings of the fourth world congress of African Linguistics, New Brunswick 2003*, eds. Akin Akinlabi and Oluseye Adesola, 211-223. Köln: Rüdiger Köppe Verlag.

Jason Kandybowicz

Kandybowicz, Jason. 2008a. *The grammar of repetition: Nupe syntax at the syntax-phonology interface*. Amsterdam: John Benjamins.

Kandybowicz, Jason. 2008b. Predicates clefting in parallel. Ms. Swarthmore College.

Kandybowicz, Jason. 2009. Embracing edges: Syntactic and phonosyntactic edge sensitivity in Nupe. *Natural Language and Linguistic Theory* 27:305-344.

Kandybowicz, Jason, and Mark C. Baker. 2003. On directionality and the structure of the verb phrase: Evidence from Nupe. *Syntax* 6:115-155.

Koopman, Hilda. 1984. *The syntax of verbs*. Dordrecht: Foris.

Koopman, Hilda. 1999. Unifying predicate cleft constructions. In Koopman 2000, 366-382.

Koopman, Hilda. 2000. *The syntax of specifiers and heads*. London and New York: Routledge.

Landau, Idan. 2006. Chain resolution in Hebrew V(P)-fronting. *Syntax* 9:32-66.

Larson, Richard, and Claire Lefebvre. 1991. Predicate clefting in Haitian Creole. In *Proceedings of the twenty-first annual meeting of the north east linguistic society*, ed. Tim Sherer, 247-262. Amherst, MA: GLSA.

Lasnik, Howard. 1981. Restricting the theory of transformations: A case study. In *Explanation in Linguistics*, eds. Norbert Hornstein and David Lightfoot, 152-173. London: Longman.

Lefebvre, Claire, and Pieter Muysken. 1988. *Mixed categories: Nominalization in Quechua*. Dordrecht: Kluwer.

Lumsden, John. 1990. The biclausal structure of Haitian Clefts. *Linguistics* 28:741-759.

Lumsden, John S., and Claire Lefebvre. 1990. Predicate cleft constructions and why they aren't what you might think. *Linguistics* 28:761-782.

Manfredi, Victor. 1993. Verb focus in the typology of Kwa/Kru and Haitian. In *Focus and grammatical relations in creole languages*, eds. Francis Byrne and Donald Winford, 3-51. Amsterdam: John Benjamins.

Matushansky, Ora. 2006. Head movement in linguistic theory. *Linguistic Inquiry* 37:69-109.

Nishiyama, Kunio, and Eun Cho. 1998. Predicate cleft constructions in Japanese and Korean: The role of dummy verbs in TP/VP preposing. *Japanese/Korean Linguistics* 7:463-479.

Nunes, Jairo. 2003. Head movement, remnant movement and phonetic realization of chains. In *Syntax at sunset 3: Head movement and syntactic theory*, ed. Anoop Mahajan, 161-177. UCLA/Potsdam Working Papers in Linguistics 10.

Nunes, Jairo. 2004. *Linearization of chains and sideward movement*. Cambridge, MA: MIT Press.

Piou, Nanie. 1982. Le clivage du prédicat. In *Syntax de l'Haitien*, eds. Claire Lefebvre, Hélèn Magloire-Holly and Nanie Piou, 122-151. Ann Arbor: Karoma Press.

Stewart, Osamuyimen Thompson. 2001. *The serial verb construction parameter*. New York: Garland.

Szabolcsi, Anna. 1994. The noun phrase. In *The syntax of Hungarian*, eds. Ferenc Kiefer and Katalin Kiss, 179-274 (Volume 27 of *Syntax and Semantics*). San Diego: Academic Press.

Vicente, Luis. 2005. Towards a unified theory of movement: an argument from Spanish predicate clefts. In *Leiden Papers in Linguistics* 23, eds. Martin Salzmann and Luis Vicente, 43-67.

Vicente, Luis. 2006. The syntax of heads and phrases: A study of verb (phrase) fronting. Doctoral dissertation, Leiden University. Utrecht: LOT Dissertation Series #154.

Vicente, Luis. 2009. An alternative to remnant movement for partial predicate fronting. *Syntax* 12:158-191.

Wh-phrasal Movement and the Adjunction Analysis of Free Relatives

Tommi Leung

United Arab Emirates University, Al-Ain

Abstract

The issue of whether head movement occurs within the narrow syntax is pursued in this paper. Our discussion centers free relatives. Some recent studies suggested that free relatives are derived by wh-head movement instead of phrasal movement, which serves as an argument against Chomsky's recent claim that head movement does not exist within syntax, instead it happens within the 'phonological branch of the syntactic derivation'. We argue, however, that the head movement approach to free relatives is misplaced and moreover based on misanalysis. Instead wh-phrasal movement is the underlying mechanism for free relatives, and the free relative clause combines with the matrix clause through adjunction. A further analysis of English and Slavic free relatives converges to this conclusion.

1.   Introduction

Since Chomsky's (1970) seminar work on wh-movement and later on the advent of the *Government and Binding Theory* (Chomsky 1981, 1982), two types of movement, *i.e.,* head movement and phrasal movement, have been established that describe the 'displacement property' of language, a long-standing yet non-trivial problem raised as early as Chomsky (1955/1975). At the outset, the two types of movement exhibit distinctive traits and are subject to distinctive conditions, such as the identity of the movement attractor, the landing site of movement, the availability of locality conditions, and certainly the identity of the moved element (*e.g.,* Matushansky 2006). The theoretical status of two types of movement receives a more rigorous and abstract evaluation from the minimalist perspective since the *Minimalist Program*. Chomsky (1995) wondered whether movement in the form of Move-Alpha (Lasnik and Saito 1992) could be reduced to Move-F(eature) in which *only* the necessary formal features (*i.e.,* necessary for convergence) is/are moved to check off the attractor feature(s), a minimalist inquiry. Chomsky concluded that feature movement, while theoretically more minimalist and therefore optimal, does not virtually exist given the consideration

of PF-convergence as the Bare Output Condition. By contrast, only pied-piping (*i.e.,* movement of syntactic categories such as words or phrases) is licensed by syntactic operations. Now the question lies on whether head movement and phrasal movement should enjoy the same theoretical status from a minimalist eye, an old question to be answered in a novel way. Chomsky claimed that head movement does not exist within the narrow syntax (i.e., at the pre-spell-out level) since it violates the Extension Condition that syntactic operations should only extend the tree in the sense that the launching site should always c-command the movement trace. Head movement instead happens at the 'phonological branch of the derivation'. The only legitimate syntactic movement is phrasal movement as a result of Copy + Re (Merge), out of EPP reason, so to speak.

The decision as to whether head movement should be ruled out as a syntactic operation is not definitive. Donati (2006) listed out examples from free relative and comparative constructions and argued against such Chomsky's claim. She pointed out that English and Italian free relatives exhibit clear cases of wh-head-movement, in which the bare wh-word moves to head the embedded free relative clause, and concluded Chomsky's assumption as untenable. By re-examining English again and some Slavic languages, this paper claims that the wh-head-movement analysis of free relatives is however erroneous. On the other hand, there is abundant evidence showing that wh-phrasal movement is at work here. In the study of free relatives, one major concern is to what extent their analysis can account for the 'matching effect'. Two competing approaches, the *Head-Account* (Bresnan and Grimshaw 1978) and the *Comp-Account* (Groos and von Riemsdijk 1981), were proposed that attempt to describe such an effect. We claim that the wh-phrasal movement approach can describe the matching effect if it is coupled by a special adjunction analysis of the free relative clause. The spirit of this analysis shares with that of the *Tree Adjoining Grammar* (TAG) that treats trees (not nodes) as the basic formative of grammar (Joshi 1985, Vijay-Shanker 1987, Frank 2002), and we argue that this approach to free relatives is conceptually and empirically supported. The details of this paper are listed in the following: Section 2 summarizes the basic facts of free relatives. Section 3 contains some previous analyzes of free relatives. Section 4 brings up further data from Slavic free relatives. Section 4 and 5 describe the problems of previous approaches to free relatives. Section 6 presents the adjunction analysis of free relatives. Section 7 is the conclusion of the paper.

2.   Basic Facts of Free Relatives

As is well known, free relatives are headless relative constructions in which the embedded wh-phrase occupies an argument or an adjunct position, depending on

the subcategorization of the matrix predicate (Bresnan and Grimshaw 1978, Groos and von Riemsdijk 1981, Bury 2003, Grosu 2002, Jacobson 1995, Caponigro 2003, van Riemsdijk 2005). To list a few examples:

(1) John liked [$_{FR}$ what Mary cooked last night].   (English)

(2) Ho   mangiato [$_{FR}$ quanto hai   preparato].   (Italian)
     have-1Sg eaten   what   have-2Sg prepared
     'I have eaten what you have prepared'

(3) Agapo [$_{FR}$ opjon   me agapa].   (Greek)
     love-1sg   whoever-Acc me loves
     'I love whoever loves me.'

(4) ég hjálpa [$_{FR}$ hverjum   ég elska].   (Icelandic)
     I help   who-Dat   I   like
     'I help who I like.'

(5) Jan czyta [$_{FR}$ co   Maria mu poleciła].   (Polish)
     Jan reads   what Maria him recommended
     'Jan reads what Maria recommended him.'

One major property of free relatives, though not attested absolutely cross-linguistically, is the 'matching effect'. There are essentially two main types of matching effect, i.e., category matching and case matching. Category matching (stated in (6)) is more attested cross-linguistically. It is shown in English examples (7) and in the schema in (8):

(6) In free relatives, the free relative clause and its head have the same categorial specification.

(7) a.   I will buy [$_{DP}$ [$_{DP}$ whatever] you want to sell].
     b.   John will be [$_{AP}$ [$_{AP}$ however tall] his father was]
     c.   I'll word my letter [$_{AdvP}$ [$_{AdvP}$ however] you word yours].
     d.   I'll put my books [$_{PP}$ [$_{PP}$ wherever] you put yours]

(8) a. NP   b. AP   c. AdvP   d. PP
          /\   /\   /\   /\
         NP S   AP S   AdvP S   PP S

In general case, violations of the matching effect are ungrammatical:[1]

---

[1] Apparently, the matching effect can sometimes be violated, especially when the wh-words are 'PP-like', e.g. where, when, and how. The followings are some notable examples (Caponigro and Pearl 2009):

(9) a.   I'll reread whatever paper John has worked on.
    b.   *I'll reread on whatever paper John has worked.

In addition to category matching, some languages exhibit the matching between the case of the wh-phrase and that of the free relative clause. For instance in German (Vogel 2001):

(10) a.   Wer   nicht stark   ist, muss klug   sein.
            who-Nom not   strong is   must clever be
            'Who is not strong must be clever'

    b.   Wer/*Wen   Gott schwach geschaffen hat, muss klug sein.
            who-Nom/whom-Acc God weak   created has   must clever be
            'Who God has created weak must be clever'

In this paper, we focus mainly on categorical matching as a typical feature of free relatives.

3.   Previous Analyses

Since free relatives were studied in the early 1970s, two major analyses were proposed that aim at capturing their special grammatical properties, i.e., the *Head-Account* (Bresnan and Grimshaw 1978) and the *Comp-Account* (Groos and von Riemsdijk 1981). The Head-Account proposed that the wh-word behaves as the head of the free relative clause that is further subcategorized by the matrix predicate, which directly describes the matching effect. On the other hand, the Comp-Account proposed that the wh-phrase is not under the head position, instead the specifier position. The major evidence comes from German extraposition. In German, starting from the basic structure in (11a), the relative clause CP can be extraposed to the sentence-final position (i.e., after the main verb) and strands the head noun (*e.g.,* 11b), whereas the extraposition of the whole DP is ungrammatical (*e.g.,* 11c):

(i)   a.   Lily adores where this very tree grows.   b.   Lily napped where this very
                                                                    tree grows.

(ii)   a.   Lily dreaded when Jack had to go.   b.   Lily cried when Jack had to go.
(iii)   a.   Lily loathes how all thieves work secretly.   b.   Lily works how all thieves
                                                                    work secretly.

We notice that the wh-gap within the free relative clause must be PP-denoting. The following free relatives are ungrammatical since the wh-gap within the free relative clause is NP-denoting:

(iv)   a.   Lily always naps where/when/how Jack despises.
    b.   Lily adores where/when/how Jack despises.

(11) a.   Der Hans hat [das Geld, das er gestohlen hat], zurückgegeben.
        The Hans has the money that he stolen   has   returned
        'Hans has returned the money that he has stolen'

     b.   Der Hans hat [das Geld t_i], zurückgegeben, [das er gestohlen hat]$_i$

     c.   *Der Hans hat t_i, zurückgegeben [das Geld, das er gestohlen hat]$_i$

However, as shown in (12), the free relative clause can be extraposed but it may not strand the wh-word, showing that the wh-word cannot function as the head of the free relative clause. The only solution is to say that the wh-word is under the specifier position:
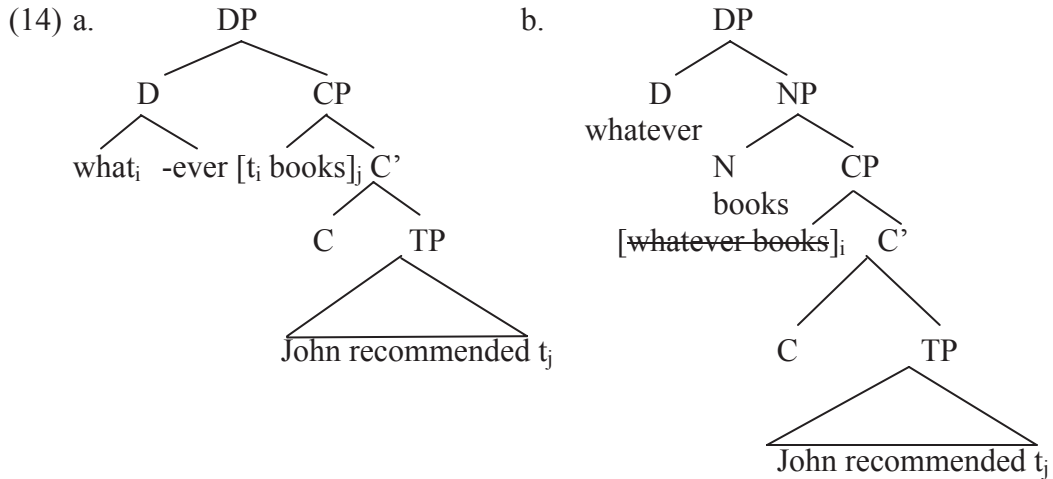
(12) a.   *Der Hans hat [was t_i] zurückgegeben [er gestohlen hat]$_i$
        The Hans has what   returned   he stolen   has

     b.   Der Hans hat t_i zurückgegeben [was er gestohlen hat]$_i$
        The Hans has   returned   what he stolen   has
        'Hans has returned what has been stolen.'

One major defect of the Comp-Account is that it misses the matching effect completely. To cope with this problem, the *Comp Accessibility Parameter* was postulated as an ad-hoc principle (Groos and von Riemsdijk 1981):

(13)   The COMP [*i.e.,* Spec-CP; TL] of a free relative is syntactically accessible to matrix rules such as subcategorization and case marking, and furthermore it is the wh-phrase in COMP, not the empty head, which is relevant for the satisfaction or non-satisfaction of the matrix requirements.

The Head-vs-Comp account of free relatives remains largely alive until recently. While it is established that free relatives that are formed by wh-words are derived by some sort of wh-movement, the identity of wh-movement is debatable. The first recent approach to free relatives is the *wh-phrasal movement* approach that originated at Kayne (1994) as a further extension of the movement approach to relative constructions. There are two versions here. The first version raised by Kayne as shown in (14a) is that the surface position of the wh-phrase is the result of overt movement. The second one is the *identity deletion approach* (14b), *i.e.,* the two free relative clauses cancel out each other before Spell-Out (Sauerland 1998, Citko 2000). Both are shown in the following[2]:

---

[2] The movement approach to relative constructions dates back to Vergnaud (1974). The proposals in (14a) and (14b) can be considered as the variants of the Head Account (see also Citko 2002).

Tommi Leung

(14) a.

```
            DP
          /    \
         D      CP
        / \    /  \
    whatᵢ -ever [tᵢ books]ⱼ C'
                         /  \
                        C    TP
                            /  \
                    John recommended tⱼ
```

b.

```
            DP
          /    \
         D      NP
    whatever   /  \
              N    CP
           books  /  \
        [w̶h̶a̶t̶e̶v̶e̶r̶ ̶b̶o̶o̶k̶s̶]ᵢ  C'
                          /  \
                         C    TP
                             /  \
                     John recommended tⱼ
```

The two approaches differ in the way they can describe the reconstruction effect (and the lack thereof). Both accurately account for Polish free relatives (Citko 2009):

(15) a.   *Idiom chunk*

? Jakiekolwiek słowaᵢ Jan rzuca tᵢ na wiatr potem go męczą.
whatever   words Jan throws on wind later him trouble
'Whatever words Jan says without thinking later come back to haunt him.'

   b.   *Variable binding*

? Zrobimy ilekolwiek portretów swojejⱼ żonyᵢ każdyⱼ król zamówił tᵢ.
  we-do   however-many portraits   self's wife   every king ordered
'We will make however many portraits of his wife every king ordered.'

   c.   *Anaphoric binding*

? Zrobimy ilekolwiek   swoichⱼ zdjećᵢ   Janⱼ zamówił tᵢ.
  we-do   however-many self's   portraits Jan ordered
'We will make however many pictures of himself Jan ordered.'

In the above cases, free relatives pattern with wh-questions with respect to the reconstruction effect, in which the wh-word is reconstructed to the based position that feeds the idiomatic interpretation and binding. In this regard, both the movement approach and the identity deletion approach are observationally adequate. However, in the following, we notice that the Principle C reconstruction is neither observed in English (16) nor in Polish (17):

(16) We will comment on [whichever pictures of Johnᵢ]heᵢ displays most prominently.

(17) a. ? Kupię na aukcji   [ilekolwiek   autoportretów Picassa$_{ik}$] on$_i$ namalował t$_k$
            buy.1sg on auction however-many autoportraits Picasso.gen he paint.p.3sg

   wokresie niebieskim.
   in period blue
   'I will buy at an auction however many autoportraits of Picasso he
   painted in the Blue period.'

   b. ? Przeczytamy [ilekolwiek książek Tołstoja$_i$ o   Rosji   w okresie
       read.1Pl   however-many books Tolstoy-gen about Russia in period

   carskim]$_k$ on$_i$ napisał t$_k$ na wygnaniu.
   tsar   he write.p.3sg on exile
   'We will read however many of Tolstoy's books about Russia in the era of
   the tsars he wrote in exile.'

These facts differ significantly with wh-questions that generally observe
Principle C reconstruction. The identity deletion seems more suitable here. The
two instances of wh-words, one generated under Spec-CP and another at the base
position, are numerated independently. The lower copy undergoes identity
deletion with the higher copy. This explains the Anti-Condition C reconstruction
since the higher copy is not the result of overt wh-movement from the base
position.

     In contrast to wh-phrasal movement, it was also proposed that
wh-movement can also apply at the head level. As is well established in the
literature, head movement differs from phrasal movement in that the former
always targets the attractor head, and the latter the Spec position. Given the
special properties of free relatives, namely the matching effect, some additional
mechanisms should be postulated to preserve the headedness of the wh-word
within the free relative clause. Bury (2003) and Donati (2006) proposed that the
wh-head can move and skip the C-head and projects its D-feature to head the DP.
Call this the *Move-and-Project* approach:

(18)
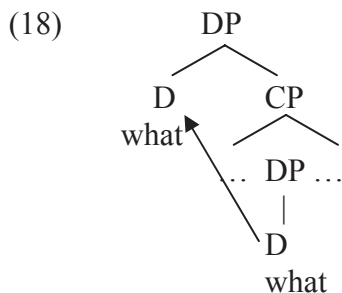
Donati showed that while English and Italian readily allow wh-phrasal movement in questions (19), only the bare wh-words can form the free relatives (20):

(19) a.   What town/what will you visit?

  b.   Quanto/quanto   pane   vorrai?
      how.much/how.much bread you.will.want
      'How much break will you want?'

(20) a.   I shall visit what/*what town you will visit.

  b.   Mangerò quanto/*quanto   pane   vorrai.
      I.will.eat how.much/how.much bread you.will.want
      'I will eat what you want.'

Based on these facts, Donati concluded that wh-head movement exists within the narrow syntax, and opposed Chomsky's (2005) assumption that it only occurs at the 'phonological branch of derivation'.

4.   Problems of Sub-extraction, Identity Deletion and the Analysis of 'What'

The complementation approach, namely the head-movement approach to free relatives, is problematic in various aspects. The first problem is that according to this approach in which the D subcategorizes for an NP or a CP and heads a DP, the D-head should not form a syntactic constituent with the head noun at the surface level (14a). In Kayne's approach, the surface position of the D-head is derived by sub-extraction of the wh-word from Spec-CP, therefore the D is separated from the NP. On the other hand, the identity deletion approach in (14b) implies that the D-head is base-generated independently of the nominal head.[3] There are examples showing that the wh-phrase is a syntactic constituent. Citko (2000) pointed out the following coordination example:

(21) I will read [whatever books] and [whatever articles] Mary recommended.

This is hardly described by (14a) or (14b). There is also another problem to the identity deletion approach. For this approach to work, we are led to conclude that

---

[3] Our proposal is not entirely contradictory to Bury's analysis of free relatives, however. Bury's view of free relatives stems from Brody's (2000) 'telescope hypothesis' in which phrase markers can be freely collapsed or canceled with each other as long as they represent the same 'treelet'. The detailed technicality has no impact on the current thesis, but the major conclusion of the telescope hypothesis is that the structural distinction between X^0 and XP disappears. Bury argued that heads can move and project their feature at the same time (same as Donati), and moreover the phrasal category can project its categorical feature. The latter claim is somehow analogous to the current proposal.

such a deletion process can only apply at the PF level instead of at the syntactic level. This is simply because the higher instance of 'whatever NP' (that do not form a syntactic constituent) is only phonologically identical to the lower instance (that form a syntactic constituent). An immediate consequence is that if identity deletion is a PF-process, the lower copy of the wh-phrase that is overtly moved to Spec-CP cannot be erased syntactically and it is still available for Spell-Out. In principle, the two copies of the wh-phrase should be available for interpretation. If the lower copy is interpreted, Principle C reconstruction would be derived, contrary to the data in (16) and (17).

Referring to (19) and (20) again, we immediately discover that the proposal of wh-head-movement to free relatives is based on impoverishment and moreover misanalysis of data. Firstly, wh-phrases are widely observed in English and Italian free relatives, especially if they are formed by the universal quantifier –*ever* that expresses a *free choice* (Jacobson 1995, Dayal 1996, Grosu 2002 and the references therein) or *speaker's ignorance* (von Fintel 2000). This can be shown in the use of 'what/whatever' (22), 'which/whichever' (23), 'how/however' (24), 'who/whoever/whoever's' (25), Italian *qualunque* 'whatever' and *qualsiasi* 'whichever' (26):[4,5]

(22) a. I will read {what/*what book/ whatever book} you suggest.

b. Do {*what/whatever} the hell you want, as long as you don't do it on paper or via broadcast.

(23) I shall visit {*which/*which town/whichever town} you will visit.

(24) a. {*How/However} many skills we possess, not all are of the same value to us.

b. You couldn't through thick smoke, {*how/however} much you wanted to.

c. He had no real intention of assaulting her, {*how/however} great the temptation which she presented to him.

d. You can write your review {*how/however} the hell you like.

(25) a. I will meet {who/whoever} you suggest.

b. I will forge {*whose/*whose signature/whoever's signature} you are willing to forge.

---

[4] Free relatives formed by -*ever* can be interpreted as concessive similar to 'no matter *wh*-' (Culicover 1999, Izvorski 2000). For instance in Culicover (1999:125):

(i)   Whatever the reason (is), I will not condone your behavior.

(ii)   Whoever the mayor (is), he can't use public funds for his own political purposes.

(iii)   I always search out fresh fish, wherever the restaurant.

(iv)   Whenever the concert, we intend to be there on time to get a good seat.

(v)   Whichever direction you take, you will find the scenery to be very pleasant.

[5] The examples come from Grosu (2003), the British National Corpus and internet search.

Tommi Leung

(26) a. Ti   siegherò   {qualunque/qualunque problema} tu   non capisca.
        to-you I.Fut.explain whatever/whatever problem you   not understand
        'I will explain to you whatever problem you do not understand.'

   b. Lui parlerebbe   con qualsiasi   ragazza gli capitasse   di incontrare.
      he would.speak with whichever girl to.him it.happened to.meet
      'He would speak with whichever girl happened to meet him.'

Secondly, the analysis of 'what' as a wh-head by Donati and Bury is unsatisfactory. It is generally assumed that English 'what' conflates quantification and restriction similar to 'which x' in the logical form (May 1985). Other wh-words such as 'where' (*i.e., which place*), 'when' (*i.e., which time*) 'who' (*i.e., which person*), 'how' (*i.e., in which way*) and 'why' (*i.e., for which reason*) can be understood in the same fashion. One can understand 'what' as a wh-phrase with the structure 'wh-NP' that contains the quantification 'wh-' and an empty nominal restriction. On the other hand, 'which' expresses quantification without any restriction. This said, it is 'which' instead of 'what' that is the typical wh-head. Morphologically, 'which' and other wh-words exhibit contrastive properties. In (27), while most wh-words are conflations of a wh-morpheme and an indefinite morpheme that can also form an anaphoric pronoun (with *th-*), 'which' represents the quantification without any restriction, and it has no anaphoric counterpart (Cheng 2001[6]):

(27)   **Interrogative/Relative pronoun**   **Demonstrative/Anaphor**
        wh-at   wh-o   th-at   th-ey
        wh-ere   wh-om   th-ere   th-em
        wh-en   wh-ich   th-en   *th-ich

Since 'which' only expresses quantification, it requires an overt nominal restriction in question formation:

(28) Which *(book) did you buy?

---

[6] One language that better demonstrates the morphological relation between wh-pronouns and demonstrative/pronouns is Polish (Citko 2000):

Wh-Pronouns		**Demonstrative/Pronouns**	
co	'what'	to	'this'
kto	'who'	ten/ta	'this'
jak	'how'	tak	'this way'
gdzie	'where'	tam	'there'
kiedy	'when'	wtedy	'then'
dlaczego	'why'	dlatego	'for this reason'
która	'which'	*	

On the other hand, 'what' can be optionally followed by a nominal restrictor or a floating quantifier 'ever' (29a) (Bresnan and Grimshaw 1978). That (29b) is ungrammatical with the indefinite 'one' suggests that the nominal restriction should be 'D-linkable' (Citko 2004).

(29) a.   What (ever) is (ever) the matter with him now?

b.   What (book/*one) did you buy?

The contrast between 'what' and 'which' extends to free relatives, in which 'whatever' can optionally be followed by a restrictor, whereas for 'whichever' an overt restrictor NP is obligatory:

(30) a.   I will buy whatever (book) you recommend.

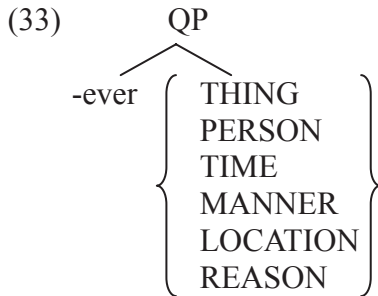b.   I will buy whichever *(book) you recommend.

Since the contrast in (30) patterns with that in (28), we have reasons to believe that English free relatives involve wh-phrasal instead of wh-head movement.[7] Further cross-linguistic examinations show that wh-phrasal movement is involved in free relatives. In Polish and Russian (Citko 2009), while it is well known that Left Branch Condition can be violated in wh-questions in which the wh-head (*e.g.,* 'which') can be extracted to the frontal position and strands the restrictor NP (*e.g.,* 31a, 32a), the condition is strictly obeyed in free relatives (*e.g.,* 31b, 32b). On the other hand, pied-piping of the whole wh-phrase is the only grammatical option (31c, 32c):

(31) a.   Którą$_i$ Jan przeczytał t$_i$ książkę?   (Polish)
            which Jan read book
            'Which book did Jan read?'

b.   *Kupię jakikolwiek$_i$ mi poleciłaś t$_i$ amochód.
             I-buy whatever   me you-recommended car
            'I will buy whatever car you recommended me.'

c.   Kupię jakikolwiek samochód$_i$ mi poleciłaś t$_i$.
            I-buy whatever   car   me recommended
            'I will buy whatever car you recommended me.'

(32) a.   Kotoruju$_i$ Ivan procital t$_i$ knigu?   (Russian)
            which   Ivan read   t$_i$ book
            'which book did Jan read?'

---

[7] It does not mean that free relatives should be treated on a par with wh-questions though. See Bresnan and Grimshaw's (1978) description of the differences between the two constructions.
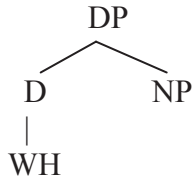
Tommi Leung

b. (*)Ja kupl'u   kakuju by ty ni porekomedoval t_i mashinu.
       I buy-will which Cond you Neg recommended   car

c. Ja kupl'u kakuju by mashinu ty   ni   porekomedoval.
       I   buy   which Cond car you Neg recommended
       'I will buy whatever car you recommended to me.'

As a result, all arguments for the wh-head movement approach to free relatives are misplaced. Now the only puzzling issue is how we account for the unacceptability of 'what NP' and 'which NP' in free relatives as in (20). While there is no perfect explanation at this moment, one can suggest that in free relatives, the presence of a restriction (regardless of its syntactic category) has to be accompanied by universal quantification expressed by –*ever*. It can be represented by the following tree (see also Culicover 1999 for a similar treatment):

(33)   QP

    -ever   ⎧ THING
             ⎪ PERSON
             ⎪ TIME
             ⎨ MANNER
             ⎪ LOCATION
             ⎩ REASON

Further incorporation combines –*ever* with the wh-morpheme that selects the QP. In the absence of a selecting wh-word, -*ever* will become *every* that is obligatorily followed by a nominal restrictor (indicated by the upper case in (33)). As a result, there exists a subcategorization relation between –*ever* and restrictions in free relatives (34b), whereas an analogous relation can be defined between a wh-word and a nominal restrictor in the context of wh-questions (34a).

(34) a. **Wh-questions**   b. **Free relatives**
             DP   DP
           /   \   /   \
          D   NP   D   QP
          | | /   \
          WH   WH   Q   NP
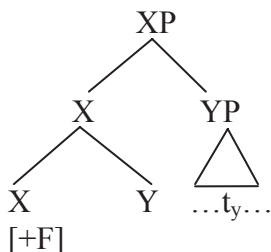   -ever   RESTRICTOR
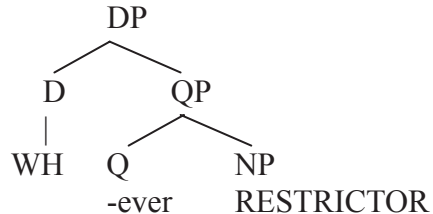
Following this line of thought, 'what' is actually 'what Q THING' in free relatives, and this analysis immediately rules out expressions such as 'what thing' or 'which thing'.
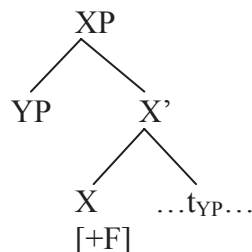
5.   Problem Two: The Concept of Wh-Head Movement

There is indeed another conceptual problem for the wh-head movement analysis to free relatives. It is generally assumed since the Minimalist Program that movement, regardless of the type of the moved categories, should only be driven by some strong formal feature(s) present at the attraction site. This spirit remains intact in the *Derivation by Phase* (Chomsky 2001), in which the Probe that searches for a Goal within the search domain and projects its uninterpretable feature and heads a phrase (also Pesetsky and Torrego 2001). As for the phrasal movement to Spec, the corresponding functional head usually embeds some strong feature(s) (*e.g.,* EPP feature, wh-feature, edge feature, etc) that require(s) its Spec position to be filled up. For head movement, some strong formal feature of the attracting head can also be postulated so that the head is adjoined by the moved element (Matushansky 2006). The two types of movement are unified at an abstract level in the sense that heads attach at the head levels, and phrases at the phrasal levels. To summarize:

(35) a.   **Head movement**   b.   **Phrasal movement**

```
         XP                              XP
        /  \                            /  \
       X    YP                        YP    X'
      / \    /\                            /  \
     X   Y  …tᵧ…                          X    …t_YP…
    [+F]                                 [+F]
```

The general assumption is that by the time movement occurs, the landing site is already constructed, such that only adjunction (as in head movement) or substitution (as in phrasal movement) is the available option. In this regard, it remains mysterious as to the type of features involved, and moreover the structural description for the Move-and-Project analysis to occur. Consider (36):

(36)
```
            YP
           /  \
          Y    XP
              /  \
             X    YP
                  /\
                …tᵧ…
```

Tommi Leung
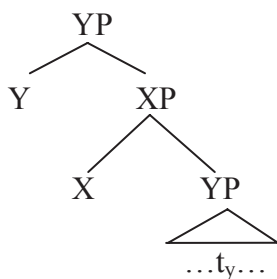
While the projection of the moved item satisfies the Extension Condition, there are at least two problems here. First, in (36) X subcategorizes for YP and moreover Y as the head of YP. However the movement of Y establishes an inversed subcategorization relation between Y and X. Second, given the *Head Movement Constraint* (Travis 1984), the movement of Y starts from its adjunction to X. How can an adjunction become a head in the second movement becomes mysterious. Syntacticians such as Bury (2003) and Citko (2007) suggested that the projection of the categories is independent of whether the projected category comes from base-generation (i.e., it comes directly from numeration) or movement. According to them, 'in principle' all the following options of Merge and Move should be seriously entertained:

(37) Merge {α, β}   (38) Move α

a.   α   b.   β   a.   β   b.   α

    α   β   α   β   α$_i$ β   α$_i$ β

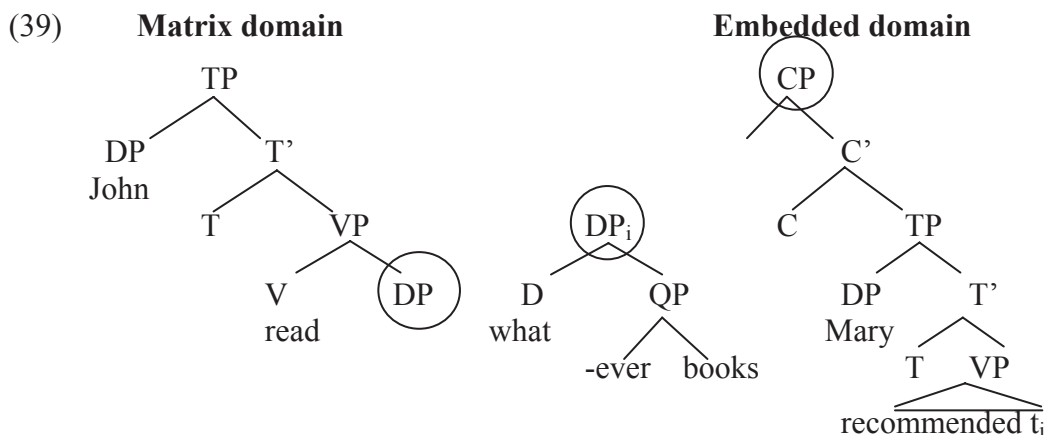   …t$_i$…   …t$_i$…
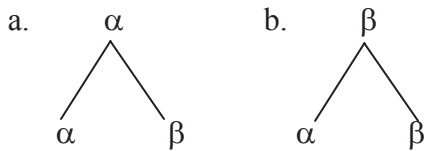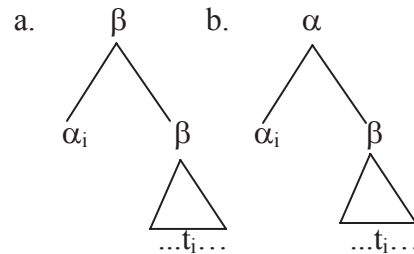
I put 'in principle' within quotations because Merge, at least in Chomsky's version, is actually asymmetric as far as the notion of labels and projections are concerned. The central issue rests upon how the label of projection is determined by the features of the merged items. For instance, Collins (2001) claimed that the projected category is the one that has one or more unsaturated features after Merge. On the other hand, the non-projected category does not have unsaturated feature left after Merge, and will be interpreted as a phrase that ceases to project. Thus there is no conceptual motivation to distinguish between (37a) and (37b), since they are simply identical to each other.

6.   The Adjunction Analysis of Free Relatives

Up to now, we have provided empirical evidence that free relatives involve wh-phrasal instead of wh-head movement. Now the central question is how we can account for the matching effect. We assume that in the derivation of free relatives such as *John read whatever books Mary recommended*, a matrix domain represented by the matrix predicate, and an embedded domain represented by the free relative clause and the moved wh-phrase, are involved:

(39)   **Matrix domain**   **Embedded domain**

The question is how to connect between the subcategorized DP in the matrix domain, the moved DP from the object trace, and the embedded CP in a way that conforms to grammatical conditions. Notice that it is not obligatory to fill up Spec-CP.[8] The moved DP 'whatever books' remains active and is free to combine with any possible target. In a sense, this is analogous to the problem raised in the *Tree Adjoining Grammar* (TAG) (Joshi 1985, Vijay-Shanker 1987, Frank 2002). One can assume, following TAG, that the moved DP substitutes the object position of the matrix domain:

(40)
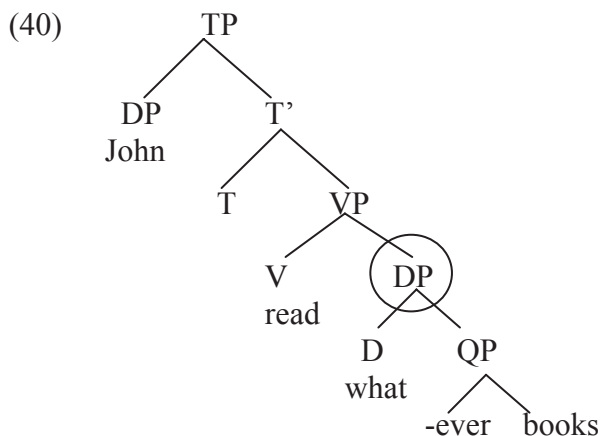
[8] Certainly the Spec-CP can be filled up by the moved DP, e.g. in interrogatives. In this case, one possible explanation is that the moved DP cannot combine with the object position of the matrix predicate, since in that case the complement of the matrix predicate is a CP that combines with the embedded clause. The only possible choice for the moved DP will be to substitute the Spec-CP position.

While the analysis can perfectly account for the matching effect, the problem lies on how the active embedded CP combines with the DP. One unsatisfactory way, originally proposed by Citko (2000, 2002), is to substitute the DP at the position of Spec-CP at the same time. Citko called this *Parallel Merge*:

(41)
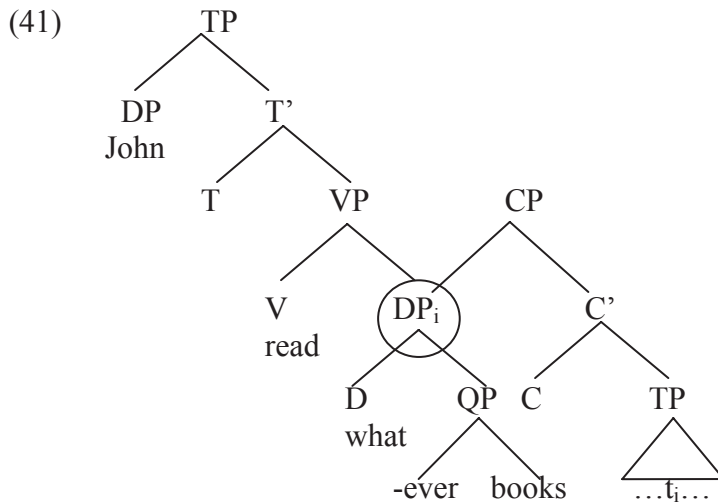
Parallel Merge needs to be supplemented by ad-hoc grammatical rules that split the trees and recombine the split trees again, otherwise the tree will be unable to linearize. On the other hand, one alternative solution is to get rid of Parallel Merge of DP as shown in (42). The moved DP functions as an *auxiliary tree* (adopting the term in TAG) in the following form:

(42)
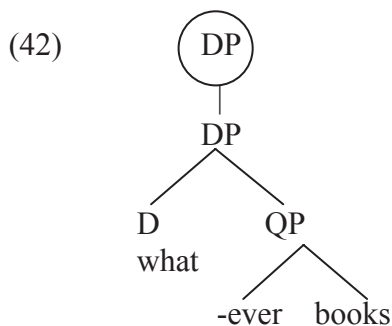
The top active node of DP substitutes the object position of the matrix predicate, while at the same time the active CP node adjoins directly to the topmost DP node, leading to the following tree:

(43)
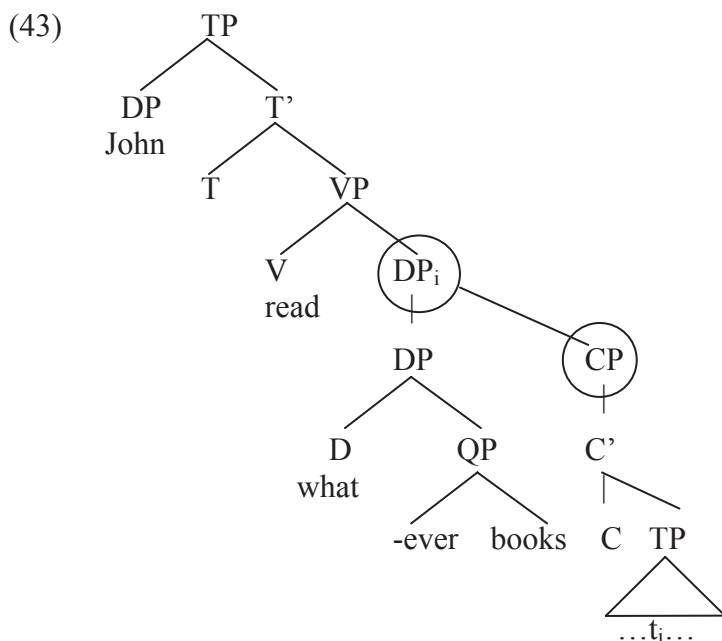
This essentially leads to an *adjunction analysis* to free relatives, contrary to the complementation approach. We notice immediately in this approach, that the embedded relative clause is an adjunction to DP. However we also understand that the relative clause is not optional. For instance:

(44) a. John read whatever *(Mary suggests).
  b. John will meet whoever *(Mary recommends).
  c. I'll put my books wherever *(you put yours).
  d. I'll do it however *(you do it).

The obligatory nature of the relative clause in free relatives does not necessarily stem from its complement nature. Rather one can understand from its semantics, namely that the relative clause contains a moved wh-trace. The trace supports the semantic interpretation of the moved DP. As a result, the free relatives are semantically vacuous without the support of the relative clause, hence its obligatoriness. The importance of the relative clause (whether it is a complement or an adjunct) is also verified in the following examples (Aoun and Li 2003):

(45) a. the Paris *(that I know).
  b. the pictures of John's *(that you lent me).

As a result, we have the following statement concerning the derivation of free relatives:

(46) In free relatives, the wh-phrase is an auxiliary tree in which the top node can be adjoined by the free relative CP.

7. Conclusion

While we remain neutral as to whether wh-head movement virtually exists, our study of free relatives concludes that the wh-head movement approach is not convincing. Instead they involve wh-phrasal movement. The main evidence comes from the reexamination of English and Slavic free relatives. In particular we rely on the adjunction analysis adopted from Tree Adjoining Grammar that consider trees as primitive objects. The special analysis is able to account for the matching effect observed in free relatives. The complement nature of the adjoined free relative clause can be accounted for by a simple semantic analysis of the wh-trace. More work should be done as to the nature of auxiliary trees in free relatives, and how auxiliary trees are connected within the computational domain.

References

Aoun, J and A. Li. 2003. *Essays on the representational and derivational nature of grammar: the diversity of wh-constructions*. Cambridge, Mass.: MIT Press.

Bresnan, J and J. Grimshaw. 1978. The syntax of free relatives in English. *Linguistic Inquiry 9*:331-391.

Bury, D. 2003. *Phrase structure and derived heads*. Doctoral dissertation, Univesity College of London.

Brody, M. 2000. Mirror theory: syntactic representation in perfect syntax. *Linguistic Inquiry* 31: 29-56.

Caponigro, I. 2003. Free relatives as DPs with a silent D and a CP complement. In *Proceedings of the Western Conference on Linguistics 2000*, ed. Vida Samiian, Fresno, CA: California State University.

Caponigro, I and L. Pearl. 2009. The nominal nature of where, when, and how: evidence from free relatives. *Linguistic Inquiry* 40.1: 155-164.

Cheng, L. L-S. 1991. *On the typology of wh-questions*. Doctoral dissertation, MIT.

Chomsky, N. 1955/1975. *The logical structure of linguistic theory*. Ms, Harvard University, Cambridge, Mass [published in Plenum, New York].

Chomsky, N. 1970. Remarks on nominalization. In *Readings in English Transformational Grammar*, eds. R. Jacobs and P. Rosenbaum, 184-221. Waltham, MA: Ginn.

Chomsky, N. 1981. *Lectures on Government and Binding*. Foris: Dordrecht.

Chomsky, N. 1982. *Some concepts and consequences of the theory of Government and*

Binding. Cambridge, Mass: MIT Press.

Chomsky, N. 1995. *The Minimalist Program*. Cambridge, Mass: MIT Press.

Chomsky, N. 2001. Derivation by phase. In *Ken Hale: a life in language*, ed. M. Kenstowicz , 1-52. Cambridge, Mass: MIT Press.

Chomsky, N. 2005. On phases. In *Foundational issues in linguistic theory*, eds. R. Freidin, C. P. Otero, and M. L. Zubizarreta, 133-166. Cambridge, Mass: MIT Press.

Citko, B. 2000. *Parallel Merge and the syntax of free relatives*. Doctoral dissertation, State University at New York, Stony Brook.

Citko, B. 2002. (Anti) reconstruction effects in free relatives: a new argument against the Comp account. *Linguistic Inquiry* 33: 507-511.

Citko, B. 2004. On headed, headless, and light-headed relatives. *Natural Language and Linguistic Theory 22*.

Citko, B. 2009. What don't wh-questions, free relatives, and correlatives have in common. In *Correlatives Cross-Linguistically*, ed. A. Liptak. Amsterdam: John Benjamins.

Collins, C. 2002. Eliminating labels. In *Derivation and explanation in the Minimalist Program*, eds. S.D. Epstein and T.D.Seely, 42-61. Oxford: Blackwell.

Culicover, P. W. 1999. *Syntactic nuts*. Oxford: Oxford University Press.

Dayal, V. 1996. *Locality in wh quantification*. Dordrecht: Kluwer Academics.

Donati, C. 2006. On wh-head movement. In *Wh-movement: moving on*, eds. L. L-S Cheng and N. Cover. Cambridge, Mass: MIT Press.

Frank, R. 2002. *Phrase structure composition and syntactic dependencies*. Cambridge, Mass: MIT Press.

Groos, A and H. von Riemsdijk. 1981. The matching effects in free relatives: a parameter of core grammar. In *Theory of markedness in Generative Grammar*, eds. Belletti et al. Scuola Normal Superiore.

Grosu, A. 2002. Strange relatives at the interface of two millennia. *Glot International* 6.6:145-167.

Izvorski, R. 2000. Free adjunct free relatives. *Proceedings of WCCFL 19*. 232-245.

Jacobson, P. 1995. On the quantificational force of English free relatives. In *Quantification in natural languages*, vol 2, eds. E. Bach., E. Jelinek., A. Kratzer, and B. H. Partee, 451-486. Dordrecht, Boston, London: Kluwer.

Joshi, A. K. 1985. How much context-sensitivity is required to provide reasonable structural descriptions: Tree adjoining grammars. In *Natural language parsing: psychological, computational and theoretical perspectives*, eds. D. Dowty, L. Karttunen, and A. Zwicky, 206-250. Cambridge: Cambridge University Press.

Kayne, R. 1994. *The antisymmetry of syntax*. Cambridge, Mass: MIT Press.

Larson, R. K. 1987. Missing prepositions and the analysis of English free relative clauses. *Linguistic Inquiry* 19:239-266.

Lasnik, H and M. Saito. 1992. *Move α*. Cambridge, Mass: MIT Press.

Matushansky, O. 2006. Head movement in linguistic theory. *Linguistic Inquiry* 37.1: 69-109.

May, R. 1985. *Logical Form*. Cambridge, Mass: MIT Press.

Pesetsky, D and E. Torrego. 2001. T-to-C movement: causes and consequences. In *Ken Hale: A life in language*, ed. M. Kenstowicz, 355-426. Cambridge, Mass: MIT Press.

Riemsdijk, H van. 2005. Free relatives: A syntactic case study. In *The Blackwell companion*

to syntax 2, eds. M. Everaert and H. van Riemsdijk, 338–382. Oxford: Blackwell.

Sauerland, U. 1998. *The meaning of chains*. Doctoral dissertation, MIT.

Travis, L. 1984. *Parameters and effects of word order variation*. Doctoral dissertation, MIT.

Vergnaud, J-R. 1974. *French relative clauses*. Doctoral dissertation, MIT.

Vijay-Shanker, K. 1987. *A study of tree adjoining grammars*. Doctoral dissertation, University of Pennsylvania.

Vogel, R. 2001. Towards an optimal typology of free relative constructions. *Proceedings of IATL 16*.

von Fintel, K. 2000. Whatever. *Proceedings of the tenth annual conference on semantics and linguistic theory* (SALT).

Word Order Typology in Graph-Theoretical Linearization
of Bare Phrase Structure[*]

Takashi Toyoshima
Kyushu Institute of Technology, Japan

Abstract

In the studies of how syntactic structures map into phonological strings of words, the Linear Correspondence Axiom (LCA) of Kayne (1994) has been most widely adopted. Yet, it is not quite compatible with the Minimalist Theory of Bare Phrase Structure (Chomsky 1994, *et seq.*). An alternative is sought in a graph-theoretical approach, proposing modifications to Kural's (2005) tree traversal linearization to solve its problems. It is shown that graph traversal is a promising approach, accounting for the three cross-linguistically common variations in word order (VSO, SVO, SOV) and the rarity of the other three logically possible variations (VOS, OVS, OSV).

1. Introduction

Since the seminal work of Kayne (1994), the issue of syntax – phonology mapping has been one of the important research topics in the minimalist program. Kayne proposed the Linear Correspondence Axiom (LCA) that maps asymmetric c-command relations in syntactic structures into linear precedence relations of terminal words. It is claimed to entail the universal Specifier-Head-Complement order, with massive reordering movements in the so-called "head-final" languages. Chomsky (1994) partially adopts LCA in his theory of Bare Phrase Structure (BPS) while keeping the phrase structures linearly unordered in the narrow syntax. One problem for BPS to fully adopt LCA is the fact that unless some non-branching projection is postulated as in Kayne (1994), no linear

---

order can be determined when the complement is a simplex terminal, since a head and its complement mutually c-command each other.

In this paper, I develop a graph-theoretical approach (Yasui 2003, Kural 2005, *inter alia*) that can effectively be applied to unordered syntactic structures of BPS. I argue that graph-theoretical linearization is a promising approach, in that typologically common VSO, SVO, and SOV word orders as the base orders can be derived from a single structure, without distinct movement for each type.

This paper is organized as follows. Section 2 surveys Kural's (2005) proposal of tree traversal linearization, briefly reviewing the tree traversals in graph theory, and in section 3, I point out two major problems in Kural (2005). In section 4, I propose modifications to Kural's tree traversal, and to the head-to-head adjunction structure in BPS, defining two types of domination relations that circumvent the directionality both in the traversal algorithm and in the branching of tree structures. Then, I demonstrate how the modified traversal linearization fares with with another common order, SOV, taking up the apparently varying word order in Vata in section 5. Finally, section 6 summarizes the paper with some concluding remarks.

2. Linearization by Tree Traversal: Kural (2005)

2.1 Tree Traversal in Graph Theory

Tree structures are not exclusive to linguistics, and they are widely used to represent a hierarchical organization of information, from a folk family tree, bibliographical cataloging, an organizational management structure, to evolutionary relationships in biology, for example, constituting a type of mathematical objects.

In set theory, a tree can be defined as a partially-ordered set $(T, <)$ such that $\forall t \in T$, the set $\{s \in T \mid s < t\}$ is well-ordered by the precedence relation $<$. In graph theory, a tree is a connected, acyclic, simple graph. As such, tree structures find a wide variety of usage in various fields, and in computer science, data structures are often modeled on them. For manipulation of data in such a tree structure, systematic ways of visiting every node, where each datum is stored, have been developed, known as tree traversal methods. They can be classified by the order of steps: performing an action on the current node (conventionally called as "visit"), or repeating the process with the subtree rooted at its child nodes. These can be effectively implemented in a recursive algorithm.

Assuming a binary tree, three basic methods can be described as follows:[1]

(1) Tree Traversals

Starting from the root, at a given node N:

   a. Preorder: Visit the node.
      i) Recursively traverse the left subtree.

   b. Inorder: Visit the node.
      ii) Recursively traverse the right subtree.

   c. Postorder: Visit the node.

Informally put, the preorder traversal first "visits" a parent node, the left child/subtree second, and then the right child/subtree last. The inorder traversal "visits" the left child/subtree first, the parent node second, and then the right child/subtree last. The postorder traversal "visits" the left child/subtree first, the right child/subtree second, and finally the parent node.

2.2 Kural's (2005) Proposal

Assuming the traditional two-dimensional X´-theoretic trees (Chomsky 1970, Jackendoff 1977), Kural (2005) proposes to apply the tree traversal methods for linearization of syntactic trees, and argues that it eliminates the need to posit structures (as specified with the Head Parameter) or derivations (with/without movement as entailed in Kayne's LCA) that vary greatly across languages.

(2)

---
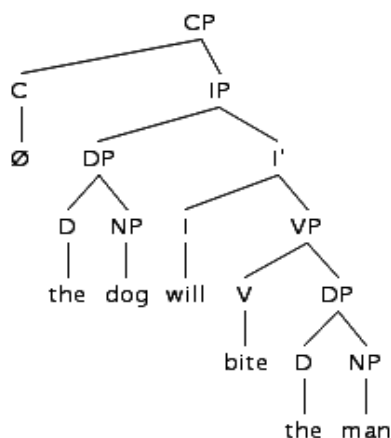
[1] Trees need not be binary, and the standard convention is to traverse children/subtrees from left to right. Also, there is another basic method known as level-order traversal, which does not seem to give rise to any linguistically significant results.

Takashi Toyoshima

Applied to a syntactic tree such as (2), the preorder traversal (1a) yields the sequence (3a), the inorder traversal (1b) the sequence (3b), and the postorder traversal (1c) the sequence (3c), respectively.

(3) a. {CP, C, Ø, IP, DP, D, *the*, NP, *dog*, I′, I, *will*, VP, V, *bite*, DP, D, *the*, NP, *man*}
   b. {Ø, C, CP, *the*, D, DP, *dog*, NP, IP, *will*, I, I′, *bite*, V, VP, *the*, D, DP, *man*, NP}
   c. {Ø, C, *the*, D, *dog*, NP, DP, *will*, I, *bite*, V, *the*, D, *man*, NP, DP, VP, I′, IP, CP}

Simply collecting the terminal words from (3a-c) produces exactly the same sequence as below, which is good enough for English but unrevealing.

(4) {Ø, *the*, *dog*, *will*, *bite*, *the*, *man*}

Kural observes, however, that there are linguistically significant patterns in (3a-c), in the relative ordering of phrasal nodes (5a-c), respectively.

(5) a. Preorder: {CP, IP, DP, NP, VP, DP, NP}
   b. Inorder: {CP, DP, NP, IP, VP, DP, NP}
   c. Postorder: {NP, DP, NP, DP, VP, IP, CP}

Replacing the phrasal nodes (5a-c) with the category symbol of their heads yields (6a-c), which coincide with the three typologically common word orders, VSO, SVO, and SOV, respectively, assuming V-to-I head movement for the VSO order in the preorder traversal (6a).

(6) a. Preorder: {C, I, D, N, V, D, N} (C)(I)SVO $\rightarrow$ (C)V(+I)S(t_V)O
   b. Inorder: {C, D, N, I, V, D, N} (C)S(I)VO
   c. Postorder: {N, D, N, D, V, I, C} SOV(I)(C)

Proposing a terminal-extraction algorithm embedded in the traversal methods (1a–c), Kural argues that tree traversal linearization can produce cross-linguistic variations in basic word order without recourse to excessive movements as in Kaynes' (1994) antisynmetry theory, that appear to be motivated only for obtaining the desired word orders.

3. Problems of Tree Traversals

Ingenious as it is, however, there are two major problems in Kural's proposal of applying the tree traversal methods for linearization of syntactic structures, one empirical and the other theoretical. The empirical problem is concerned with a certain type of movement, *wh*-movement, in particular, as in (7).

(7)
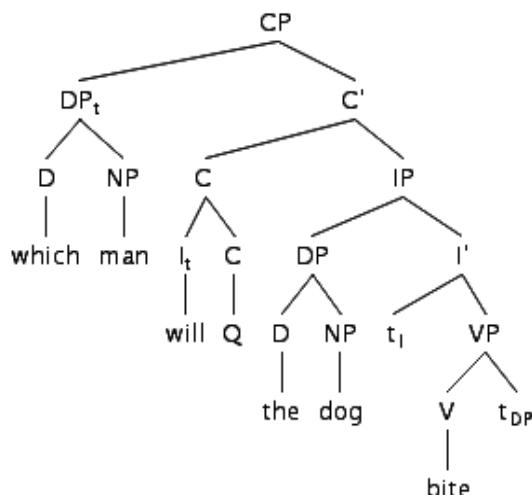
The preorder traversal, *i.e.*, for VSO languages, yields the sequence (8a), and its categorial reduction (8b).

(8) a. {CP, DP, D, *which*, NP, *man*, C′, [C, I, *will*, C, Q], IP, DP, D, *the*, NP, *dog*, I′, t_{will}, VP, V, *bite*, t_{wh}}

   b. {C, D, N, I, D, N, V}

An English gloss version will be (9) in which the moved *wh*-phrase follows the complementizer, which does not seem to be attested in any VSO languages.[2]

(9) {*will + Q which man the dog* (t_{will}) *bite* (t_{wh})}

The theoretical problem has to do with the directionality (left/right subtrees) referred to in the traversal algorithms (1a-c), and the assumption that syntactic structures are two-dimensional X′-theoretic trees which are, in graph-theoretical terms, rooted, directed, labeled, ordered, plane trees.

   Graph-theoretically, trees can be unordered, meaning an ordering is not specified for the children of each node (vertex), and the reference to the left/right subtrees is only possible in an ordered tree. A tree is necessarily a *planar graph*; that is, it *can be* drawn on a Euclidean plane without any crossing of branches

---

[2] Unlike English embedded topicalization, Irish, a VSO language, has preposing of adverbials, heavy NPs, and emphatic constituents, the last of which McCloskey (1996) calls narrative fronting. All of them front the material above the complementizer.

(edges), but need not be a *plane graph*, meaning it need not be cast into a two-dimensional plane in reality.

In the Minimalist Theory of Bare Phrase Structure (Chomsky 1994), phrase structures are incrementally built up step by step, by the operations Merge and Move, and there are no X´-theoretic schema of any sort that make phrase structures to be ordered plane trees. Rather, phrase structures are taken to be unordered without any specific branching direction. It is further suggested that adjunct modifiers may belong to a distinct dimension from the ones for the core phrase structure, implying dimensions higher than three.

Contra Kayne (1994), there is no non-branching projection, and in principle, multiple specifiers are allowed in Bare Phrase Structure (BPS). Projections are labeled with a lexical item, of which is the head of the projection, and the levels of projections are taken to be relational properties, derivationally determined by the structural configuration at a given stage in the derivation. If need be, they can be annotated with [± maximal, ± minimal] features, but they are to be understood as theoretical diacritics since they are not inherent to the structures and subject to changes in the course of derivation.

The concern for the directionality in the traversal algorithms and in branching of trees is acknowledged and discussed in Kural (2005, pp.385–6), but it falls short of any further insights. As Kural points out, it is just conventional to traverse a tree from left to right, and nothing theoretically prohibits traversals from starting from right to left. Kural illustrates that reversing the traversal order does not affect the result if the branching in the tree is also reversed.

The preorder, inorder, and postorder traversals of the following tree (10) yield the respective sequences in (11a–c) with the conventional left-to-right traversal.

(10)

```
        A
       / \
      B   C
         / \
        D   E
```
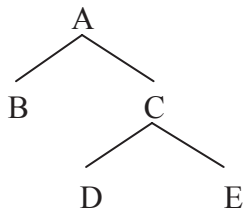
(11) a. Preorder: {A, B, C, D, E}
    b. Inorder: {B, A, D, C, E}
    c. Postorder: {B, D, E, C, A}

The tree in (12) is the mirror image of (10), branching right-to-left. Applying the reverse traversals from right to left yield exactly the same sequences as in (13a–c).

(12)
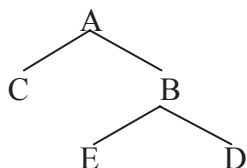
(13) a. Reverse Preorder: {A, B, C, D, E}

b. Reverse Inorder: {B, A, D, C, E}

c. Reverse Postorder: {B, D, E, C, A}

Kural claims that once grammar set the branching direction of the trees, it "feeds into the traversal algorithms." That is, if the branching direction of the trees is set as left to right, then the order in the traversal algorithms will also be set as left to right. However, it is not at all clear why the branching direction of the tree should covary with the traversal direction of the algorithm. Graph-theoretically, the branching direction of trees and the traversal direction of the algorithms are logically independent, and nothing seems to impose the same directionality in both. Thus, Kural further argues paradoxically that the reportedly rare VOS, OVS, and OSV orders can readily be derived by countering the traversal direction against the branching direction, without demonstrating the actual processes of traversal linearization, leaving their verifications to the reader.

Let us see, then, how they work out, with the same tree (2), repeated below as (14), applying the reverse traversals (15).

(14)
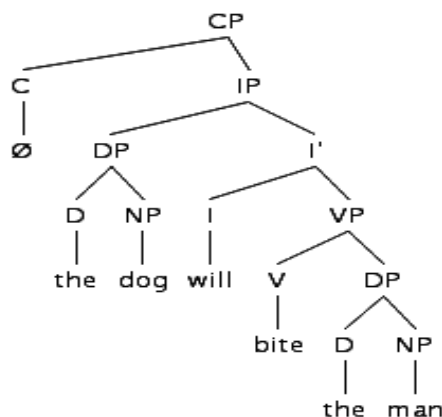
(15) Reverse Traversals
Starting from the root, at a given node N:

a. Preorder: Visit the node.
   i) Recursively traverse the **right** subtree.

b. Inorder: Visit the node.
   ii) Recursively traverse the **left** subtree.

c. Postorder: Visit the node.

The respective traversals yield the following sequence.

(16) a. {CP, IP, I′, VP, DP, NP, *man*, D, *the*, V, *bite*, I, *will*, DP, NP, *dog*, D, *the*, C, Ø}

b. {*man*, NP, DP, *the*, D, VP, *bite*, V, I′, *will*, I, IP, *dog*, NP, DP, *the*, D, CP, C, Ø}

c. {*man*, NP, *the*, D, DP, *bite*, V, VP, *will*, I, I′, *dog*, NP, *the*, D, DP, IP, Ø, C, CP}

Extracting only the phrasal nodes produces the following sequences, respectively.

(17) a. Reverse Preorder: {CP, IP, VP, DP, NP, DP, NP}
b. Reverse Inorder: {NP, DP, VP, IP, NP, DP, CP}
c. Reverse Postorder: {NP, DP, VP, NP, DP, IP, CP}

And their categorial reductions are the following:

(18) a. Reverse Preorder: {C, I, V, D, N, D, N}   (C)(I)VOS
b. Reverse Inorder: {N, D, V, I, N, D, C}   OV(I)S(C)
c. Reverse Postorder: {N, D, V, N, D, I, C}   OVS(I)(C)→ O(t_V)SV(+I)(C)

Assuming the V-to-I movement, the Reverse Postorder (18c) instantiates the OSV word order, one of the three rare cases.

Kural seems to be taking the fact that these typologically rare word orders can be derived from the same single structure, to be one of the theoretical advantages over the traditional analyses of word-order variations in terms of the Head Parameter or the more recent ones in terms of Kayne's (1994) LCA. Yet, it strikes me that it is not at all desirable. Do we want these rare word orders of

VOS, OVS, and OSV to be the base orders stemming from the same single structure? That is, without any relevant movement?[3]

My answer is negative. If VOS, OVS, and OSV orders really rare, even if they exist, they ought to be derived by some exotic movements that the majority of the world's languages do not employ; hence, the rare orders. They should not be equally derivable from the same single structure as the other three cross-linguistically common word orders, VSO, SVO, and SOV.

4. BPS Traversals: Proposal

4.1 Not Left or Right

Given the empirical and theoretical problems as we have seen in the previous section, I propose several modifications to Kural's tree traversal linearization.

The first is the adoption of BPS as syntactic trees to be traversed, which are unordered multi-dimensional trees. In order to refer to unordered children/subtrees, we need some ways of distinguishing them for reference. Thus, I define the two distinct relations in domination and the non-distinctness of labels as follows:

(19) *Domination Relations*

    a. A child/subtree is *consanguineous* if its label is non-distinct from the one of its parent.

    b. A child/subtree is *adopted* if its label is distinct from the one of its parent.

(20) *Label Non-Distinctness*

    Labels are *non-distinct* iff they are of the projections of the same token of the same lexical item.

Instead of left or right, we refer to the children/subtrees as *consanguineous* or *adopted*, dispensing with the ordering direction. Assuming the binary branching of BPS, the *consanguineous* children/subtrees inherit the same label of their parent node, modulo the projection level.[4] That is, a *consanguineous* relation picks out a head projection, while the *adopted* relations are of a specifier, a complement, or an adjunct.

---

[3] Modulo the V-to-I movement for the OSV order.

[4] The inheritance of the label is from the perspective of the dominance relations in a tree. From the perspective of derivation, the parent node inherits the label from a child/subtree. The projection levels are always distinct, but in practice, they are not usually annotated. They are determined in the structural configuration at a given point in derivation, as will be discussed in the following subsection.

Takashi Toyoshima

Then, I modify the traversal algorithms as below, making the action performed by "visit" more explicit.

(21) **BPS Traversals**

Starting from the root, at a given node N:

a. Preorder: Spell-Out the maximal label
   i) If a child is consanguineous but childless, traverse that child.
      Otherwise, traverse its adopted subtree recursively.

b. Inorder: Spell-Out the maximal label
   ii) Recursively traverse its consanguineous subtree.

c. Postorder: Spell-Out the maximal label

By the maximal label, I mean the label of a maximal projection, which can be readily read off in the structural configuration as discussed in the next subsection. The projection levels do not have to be indicated on their labels. If the parent of a given node bears a distinct label, the label of that node is maximal; otherwise, non-maximal. To put another way, if the parent is **consanguineous**, the given node bears non-maximal label; otherwise, maximal.

4.2 The First Try

Now, let us see how these modifications fare with the problematic case of *wh*-movement in (7). The BPS representation of the two-dimensional X´-theoretic tree (7) would be something like the following:

(22)

Word Order Typology in Graph-Theoretical Linearization of Bare Phrase Structure
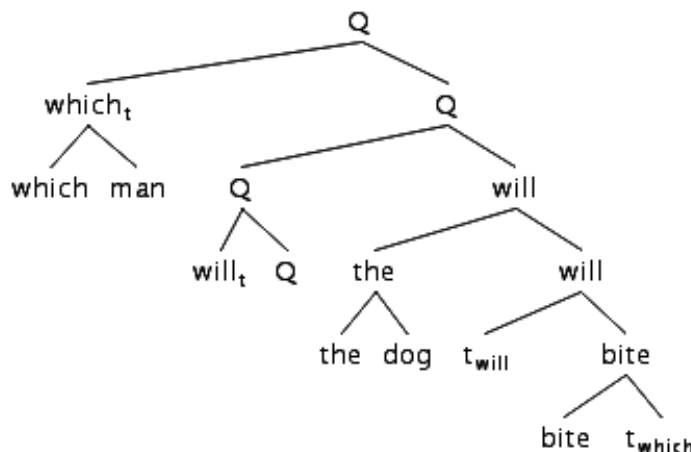
Here, the **bold italic** indicates a maximal projection, the *plain italic* a minimal projection, and the **plain bold** an intermediate projection, *i.e.*, [– maximal, – minimal], for the ease of exposition. The BPS Traversal in preorder (21a) of the BPS (22) yields the following (partial) sequence:

(23)   {***Q***, ***which***$_t$, ***man***, *which*, **Q**, … }

Starting from the root ***Q*** of the entire structure, the child node **Q** is consanguineous but with children. Thus, the algorithm traverses the adopted subtree of ***Q***, which is the subtree rooted at ***which***. Recursively starting from the node ***which*** as the root, the child *which* is consanguineous and childless, so the algorithm traverses it next. Then, the adopted child ***man*** is traversed, and then the algorithm back-tracks to the parent ***Q*** and traverse its consanguineous child **Q**.

The next step is to traverse a child of **Q**, but both of its children are not childless: subtrees rooted at *Q* and at ***will***. Then, the algorithm (21i) dictates that the adopted subtree to be traversed in preorder before the consanguineous subtree (21ii). Thus, the traversal proceeds on to the adopted subtree rooted at ***will***, and the final sequence produced will be the following:

(24)   {***Q***, ***which***$_t$, ***man***, *which*, **Q**, ***will***, ***the***, ***dog***, *the*, **will**, t_{will}, ***bite***, *bite*, t_{which}, *Q*, *Q*, ***will***$_t$}

Note, in passing, that the I-to-C moved head ***will*** is a maximal projection as well as a minimal projection; it does not project any further at the head-adjoined position nor is a projection of any other category. Thus, by definition, it is [+ maximal, + minimal], a *bona fide* feature specification in BPS.

(25)   Given a phrase marker, a category that does not project any further is a maximal projection XP, and one that is not a projection at all is a minimal projection X^0; any other is an X′, invisible at the interface and for computation.   (Chomsky 1994: 10)

Categorial reduction of (24) by the action step "Spell-Out the maximal label" in the algorithm (21a) yields the following sequence:[5]

(26)   {***Q***, ***which***$_t$, ***man***, ***will***, ***the***, ***dog***, ($t_{will,}$) ***bite***, ($t_{which,}$) ***will***$_t$}

The moved *wh*-phrase ***which man*** still follows the question morpheme ***Q***, but ***Q*** does not have any phonetic form and hence it is not pronounced, so the same

---

[5] Hereafter, I indicate the trace position in the categorial reduction for the ease of understanding their relative orders with the spelled-out maximal labels.

problem with Kural's proposal does not arise. The trace copies t_{will} and t_{which} are not pronounced, either.

Yet, a new problem arises in (26); there are two **will**'s. As the trace copies are not pronounced, it is not at all unnatural to delete excessive copies phonetically; the problem in this case is which copy to delete. It seems natural to think that all the non-first copies are to be deleted and it is the first copy that is pronounced. And as it is, the first **will** is correctly placed between the *wh*-moved **which man** and the subject **the dog** in (26). However, the first **will** is the maximal label of the IP projection, and it does not reflect the I-to-C movement of the "subject-auxiliary inversion." The I-to-C moved **will**$_t$ is the second copy that appears in the final position in (26). This is clearly not a satisfactory account.

4.3 Head-to-Head Fusion, Not Adjunction

The heart of the problem lies in the segmented head-to-head adjunction structure. Consider the following configuration, depicted in the traditional X´-theoretic tree labeling for the ease of discussion.

As Gaertner (1995) points out, the upper segment X_1 should count as [– maximal, – minimal] projection, since it "projects" to X´ and further to XP while it "is a projection" of the lower segment X_2. Then, the whole head-adjoined complex head $X_1 = [H + X_2]$ should be "invisible" to any further operations, just as X´ categories are, as defined in the theory of BPS; see (25). The result would be that no head-adjoined complex head can further be moved, contrary to the standard assumption of successive cyclic head-to-head adjunction, such as V-to-I-to-C movement, for example.

(27)

XP
YP   X'
      X_1   ZP
      H   X_2
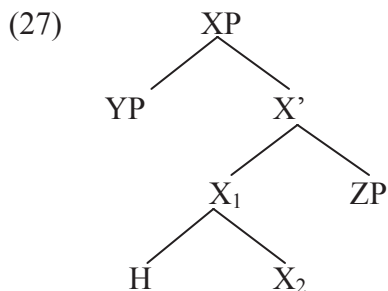
In a set-theoretic notation, Chomsky (1994) distinguishes segment projections from categorial projections by labeling. For categorial projections, the label is the lexical item that heads the projection. The label of a segment projection is an

ordered-pair of the lexical item that heads the projection. Thus, in the following tree configuration, with α being the head of the structure, β the complement, δ a specifier, and ζ an adjunct, ε_1 and ε_2 are segments.

(28)

$$\varepsilon_1 = \{\langle\alpha, \alpha\rangle, \{\{\{\alpha, \beta\}, \delta\}, \zeta\}\} = \{\langle\alpha, \alpha\rangle, \{\{\gamma, \delta\}, \zeta\}\} = \{\langle\alpha, \alpha\rangle, \{\varepsilon2, \zeta\}\}$$

$$\zeta$$

$$\varepsilon_2 = \{\langle\alpha, \alpha\rangle, \{\{\alpha, \beta\}, \delta\}\} = \{\langle\alpha, \alpha\rangle, \{\gamma, \delta\}\}$$

$$\delta$$

$$\gamma = \{\alpha, \{\alpha, \beta\}\}$$

$$\alpha \qquad \beta$$
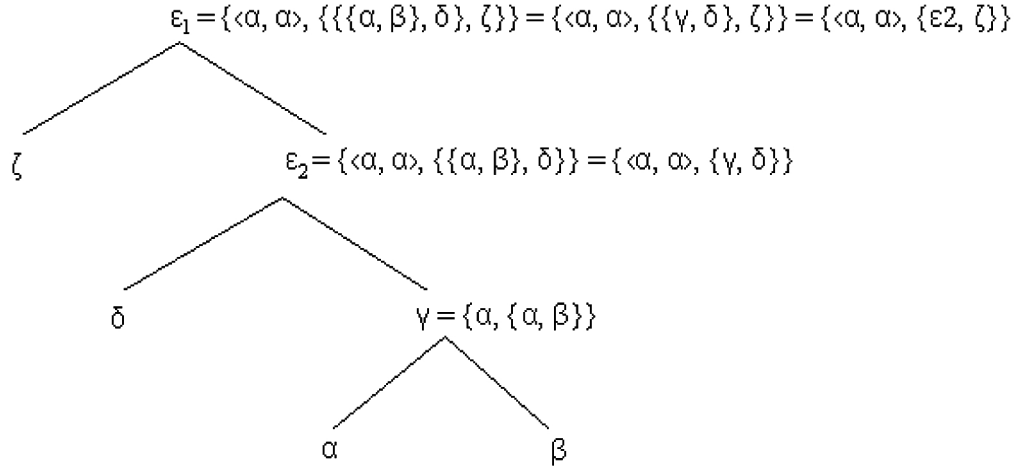
This is a segmented structure with the phrasal adjunct ζ, and its set-theoretic notation is sufficiently complicated already. Yet, when we consider a segmented structure with head adjunction, it would be more complicated as in the following:

(29)

$$\varepsilon = \{?, \{\{\{\iota, \alpha\}, \beta\}, \delta\}\} = \{?, \{\{\alpha, \beta\}, \delta\}\} = \{?, \{\gamma, \delta\}\}$$

$$\delta$$

$$\gamma = \{?, \{\{\iota, \alpha\}, \beta\}\} = \{?, \{\alpha, \beta\}\}$$

$$\{\langle\alpha, \alpha\rangle, \{\iota, \alpha\}\} = \alpha$$

$$\beta$$

$$\iota \qquad \alpha$$

Takashi Toyoshima
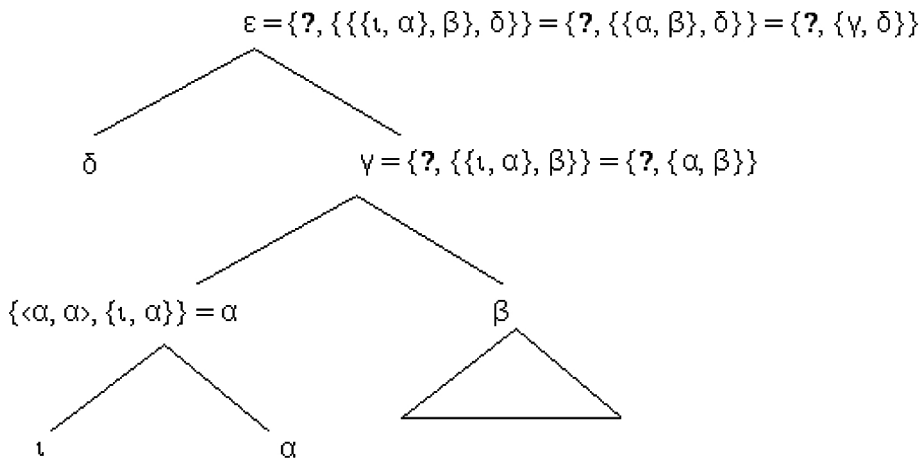
Here, α is the head of the maximal projection ε, to which another head ι is adjoined and segmented α, β is the complement, and δ is the specifier. It is not clear how γ and ε should be labeled; should they be labeled as ‹α, α›, since they are projecting from the upper segment of α whose label is ‹α, α›? Or, should they be labeled simply as α, since their head is α and ‹α, α› is not their head but the label of only the upper segment of their head α?

Chomsky (1994) does not elucidate this point, but an ordered-pair of an identical element does not seem to make much sense, be it for a segment projection or for a segmented head. Set-theoretically, an ordered-pair in the so-called Kuratowski's definition is the following:

(30)   ‹α, β› := {{α}, {α, β}}

The ordered-pair ‹α, α› is therefore equivalent to the set {{α}, {α, α}}. As any repetition of an identical element in a set is irrelevant and redundant, it is equivalent to the singleton set of that element. That is, the set {α, α} is equivalent to the singleton set {α}. Thus, the set {{α}, {α, α}} is equivalent to the set {{α}, {α}}, which in turn is equivalent to {{α}}. In short, ‹α, α› = {{α}}.

Meanwhile, since the empty set is a subset of any set, the set {{α}} is equivalent to {{∅, α}} = {∅, {∅, α}}, which conforms to the labeled set notation of BPS. It can be graphically represented as a piece of BPS:

(31)

$$\eta = \{\varnothing, \{\varnothing, \alpha\}\} = \{\{\varnothing, \alpha\}\} = \{\{\alpha\}\} = \{\{\alpha\}, \{\alpha\}\} = \{\{\alpha\}, \{\alpha, \alpha\}\} = \langle\alpha, \alpha\rangle$$

∅   α
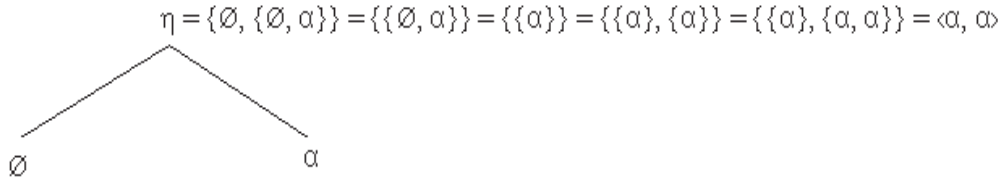
Here, the empty set ∅ is the head of the structure η, whose label is thus ∅. Since the label of a structure is determined by the head of the structure, the empty set ∅ ends up being the head of η. Furthermore, since η is the label ‹α, α› of the upper segment of α in (29), the head of the structure ε should be η whose head is ∅. Therefore, we may say that the label of the structure ε is ∅. This state of affairs can be represented in the BPS tree by substituting ‹α, α› in (29) with the empty set ∅ as follows:

(32)

$$\epsilon = \{\varnothing, \{\{\{\iota, \alpha\}, \beta\}, \delta\}\} = \{\varnothing, \{\{\alpha, \beta\}, \delta\}\} = \{\varnothing, \{\gamma, \delta\}\}$$

δ

$$\gamma = \{\varnothing, \{\{\iota, \alpha\}, \beta\}\} = \{\varnothing, \{\alpha, \beta\}\}$$

$$\{\varnothing, \{\iota, \alpha\}\} = \alpha$$

β

ι

α
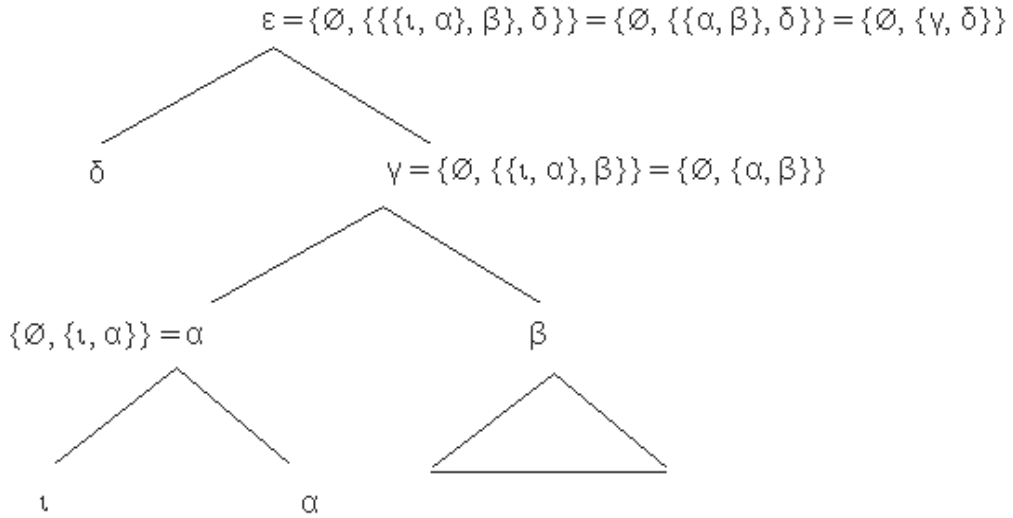
Alternatively, the singleton set $\{\alpha\}$, not the lexical item α itself, could be the head of the structure, which would look like the following:

(33)

$$\epsilon = \{\{\alpha\}, \{\{\{\iota, \alpha\}, \beta\}, \delta\}\} = \{\{\alpha\}, \{\{\alpha, \beta\}, \delta\}\} = \{\{\alpha\}, \{\gamma, \delta\}\}$$

δ

$$\gamma = \{\{\alpha\}, \{\{\iota, \alpha\}, \beta\}\} = \{\{\alpha\}, \{\alpha, \beta\}\} = \langle\alpha, \beta\rangle$$

$$\{\{\alpha\}, \{\iota, \alpha\}\} = \alpha = \langle\alpha, \iota\rangle$$

β

ι

α
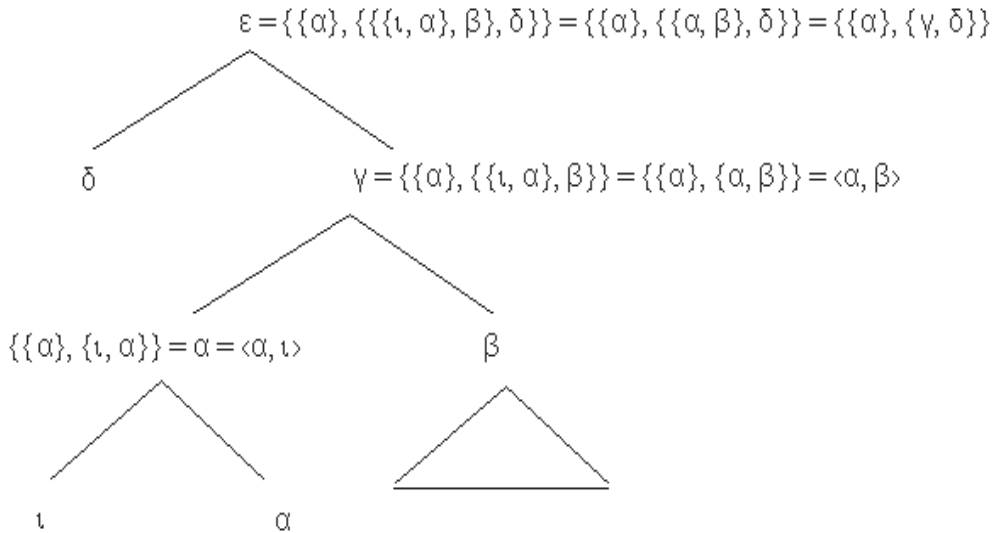
In (32), the structure ϵ ends up being exocentric, with its projections having the label $\varnothing$, distinct from the head α of the structure. In (33), in addition to the

Takashi Toyoshima

exocentricity of the structure ε (α ≠ {α}), the upper segment of α and the constituent γ are ordered-pairs, not labeled sets as defined in BPS. This state of affairs contradicts the basic tenet of the BPS theory that the phrase structure is an unordered endocentric tree.

Given these complications, I propose to modify the head-adjoined structure as follows:

(34)

$$\varepsilon = \{\alpha, \{\{\{\iota + \alpha\}, \beta\}, \delta\}\} = \{\alpha, \{\gamma, \delta\}\}$$

δ

$$\gamma = \{\alpha, \{\{\iota + \alpha\}, \beta\}\}$$

ι + α

β
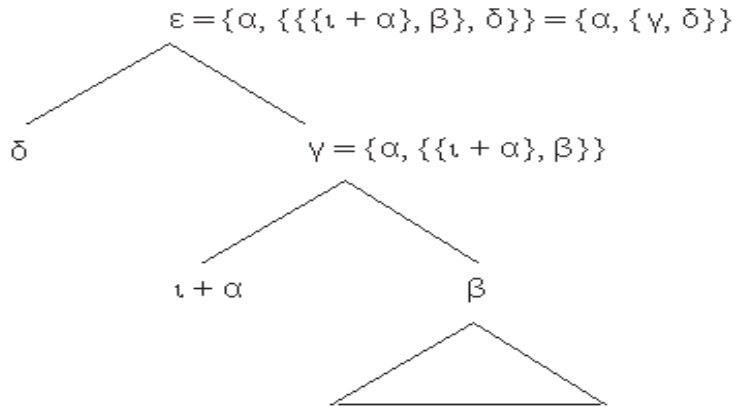
The head-to-head movement structure does not branch out, but rather fuses as one morphological unit, with the hosting head keeping its consanguineous status as non-distinct in the structure. Instead of an ordered-pair of the head as the label (35a), the complex head has the labeled set representation (35a).

(35) a. {⟨α, α⟩, {ι, α}}

   b. {α, {ι + α}}

That is, head-to-head movement is not head-to-head adjunction but rather head-to-head fusion, so to speak. The hosting head remains as the consanguineous head of the projection, and the moved head becomes a maximal "projection," just as adjoining head became at the head-adjoined landing-site; they do not project any further from there.

4.4 Retry

Given these modifications I propose in the preceding sections, the structure (22) in question would look like the following:

(36)
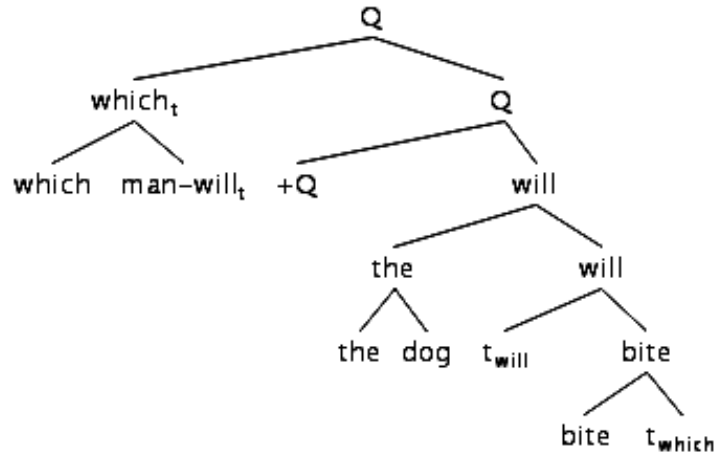
The preorder in BPS Traversal (21a) in the head-fused BPS (36) produces the same sequence (23) up to the node **Q**, repeated as (37) below.

(37)   {$\textbf{\textit{Q}}$, $\textit{\textbf{which}}_t$, $\textit{\textbf{man}}$, $\textit{which}$, **Q**, … }

The next step is to traverse a child of **Q**. In the head-adjoined BPS (22), the both children of **Q** were not childless, so the algorithm (21i) previously chose the adopted subtree rooted at **will** to be traversed, yielding an undesirable sequence.

In the head-fused BPS (36), the children of **Q** are the head-fused [**will** + Q] and the subtree rooted at **will**. As stipulated, the head-fused [**will** + Q] is still consanguineous to its parent **Q** as it includes Q and is childless as it does not branch out. Thus, the traversal proceeds on to this head-fused [**will** + Q] node first (21i), and then goes on to the subtree rooted at **will**. The traversal from the node rooted at **will** proceeds as before, and the following sequence will be produced.

(38)   {$\textbf{\textit{Q}}$, $\textit{\textbf{which}}_t$, $\textit{\textbf{man}}$, $\textit{which}$, **Q**, $\textit{will}_t$ + Q, **will**, $\textit{\textbf{the}}$, $\textit{\textbf{dog}}$, $\textit{the}$, **will**, t_{will}, $\textit{\textbf{bite}}$, $\textit{bite}$, t_{which}}

Categorial reduction of (38) by the action step "Spell-Out the maximal label" in the algorithm (21a) yields the following sequence:

(39)   {$\textbf{\textit{Q}}$, $\textit{\textbf{which}}_t$, $\textit{\textbf{man}}$, $\textit{\textbf{will}}_t$, **will**, ($t_{will,}$) $\textit{\textbf{the}}$, $\textit{\textbf{dog}}$, $\textit{\textbf{bite}}$(, t_{which})}

In this reduction, the first $\textit{\textbf{will}}_t$ was originally the head of the IP projection, fused with Q by head-to-head movement and became maximal. The second **will** registers the otherwise *in-situ* spell-out position of the head *qua* the IP projection.

Takashi Toyoshima

Now, we can safely delete this second ***will*** *in-situ*, properly reflecting the I-to-C movement chain whose head is righteously spelled-out and whose tail is deleted as usual.

5. Word Order Typology

Given these modifications I have proposed, the inorder and postorder BPS Traversals in the head-fused BPS (36) yield the following sequences, respectively, after categorial reduction and deletion.

(40) a. {***which**_t_*, ***man***, ***Q***, ***will**_t_*, ***the***, ***dog***, ~~***will***~~, (t_{will},) ***bite***(, t_{which})}

    b. {***man***, ***which**_t_*, ***will**_t_*, ***dog***, ***the***, (t_{will},) (t_{which},) ***bite***, ~~***will***~~, ***Q***}

A plain English *wh*-question is instantiated in (40a) as the object *wh*-question in SVO languages: $[O_{wh}(C+I)S(I_t)V(Ot_{wh})]$, properly reflecting the head-to-head movement chain ‹***will**_t_*, (~~***will***~~, t_{will})›, and the *wh*-movement chain ‹***which**_t_* ***man***, t_{which}›.

The postorder traversal is for SOV languages as in Kural's (2005) proposal, and (40b) may be instantiating the object *wh*-question in Vata as reported in Koopman (1983). Categorially, (40b) has the following sequence:

(41)   {N, D$_{wh}$, I$_t$, N, D, ($t_\mathbf{I}$,) (N, D$_{wh}$,) V, ~~I~~, C$_Q$}

The first **N − D$_{wh}$** sequence is the *wh*-moved object and the second **N − D** sequence is the subject. The first **I$_t$** is the moved head of the IP projection while the second ~~I~~ reflects the *in-situ* position of the moved **I$_t$**. Were there no I-to-C movement, the sequence would have been the following:

(42)   {N, D$_{wh}$, N, D, ($t_\mathbf{I}$,) (N, D$_{wh}$,) V, I, C$_Q$}

Koopman (1983) cites a simple object *wh*-question in Vata such as the following:

(43)   àlÓ Kòfĩ yÉ̀ *t*_wh_ *t*_v_ yé lá
        who Kofi saw   PRT Q
        'Who did Kofi see?'

Assuming that the particle *yé* is an auxiliary and the verb *yÉ̀* 'saw' moves to I, (43) does appear to match the categorial sequence (42), instantiating the object *wh*-question in SOV languages without "subject-auxiliary inversion." A BPS representation with an English gloss will be (44), with the branching directions

appropriated for SOV languages in the manner of the more familiar X´-theoretic representation for the ease of illustration.

(44)
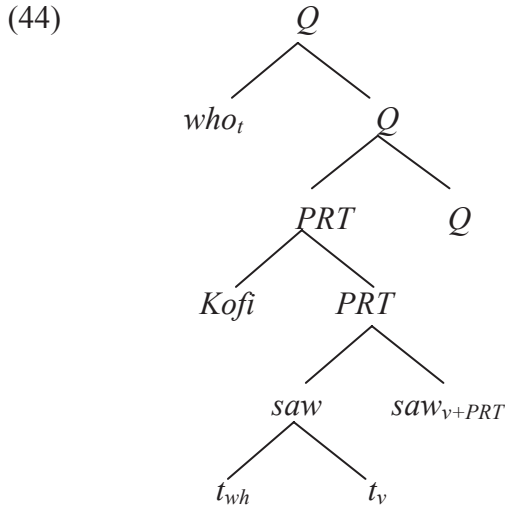
(45) {*who*, *Q*, *Kofi*, *saw* + *PRT*, (*t*$_v$,) (*t*$_{wh}$,) ~~*saw*~~, **PRT**, **PRT**, ***Q***}

(46) {*who*, *Kofi*, *saw*, (*t*$_v$,) (*t*$_{wh}$,) ~~*saw*~~, **PRT**, ***Q***}

(47) { àlÓ, Kòfĩ, yÉ̀, (*t*$_v$,) (*t*$_{wh}$,) ~~yÉ̀~~, yé, lá}

The postorder traversal in BPS (44) yields the sequence (45) whose categorial reduction is (46), which in turn is converted with the Vata lexical items in (47).

Now, compare (47) with the original Vata object *wh*-question (43), where the relative order between the trace of the moved verb (*t*$_v$,) and the one of the *wh*-object (*t*$_{wh}$,) is reversed. As the traces are not pronounced, they do not technically cause apparent problems as they are, but they beg a theoretical question whether this is an adequate analysis. The fact of the matter is that it is indeed correct, in that the trace position of a moved head is "spelled-out" in the *in-situ* position, represented with its maximal label. Therefore, in BPS Traversal, the traces of head movement are the duplicate copies in the yielded sequence before categorial reduction, which were deleted when spelled out. Thus, in (47), the crossed-out verb ~~yÉ̀~~ 'saw' signifies the *in-situ* position of the verb with its maximal label, whereas the trace *t*$_v$ marks the trace position of the terminal verb in the traditional X´-theoretic sense, which I retained for the purpose of comparison (see fn. 5). In other words, the crossed-out maximal labels are the traces in BPS Traversal.

Takashi Toyoshima

Notwithstanding the successful account of the object *wh*-question, Vata exhibits a mixed ordering, the SVO order in simple tenses (48a) and the SAuxOV order in complex tenses (48b).

(48) a. à lì saká   b. à lā saká lī
       we ate rice   we have rice   eaten
       'We ate rice.'   'We have eaten rice.'

Koopman argues that in simplex tenses, the verb raises to I from the underlying S(I)OV structure, where the perfect auxiliary resides. If so, (48a, b) would have the respective BPS representations, as in (49a, b), shown in English gloss, omitting irrelevant details.

(49) a.

(49) b.
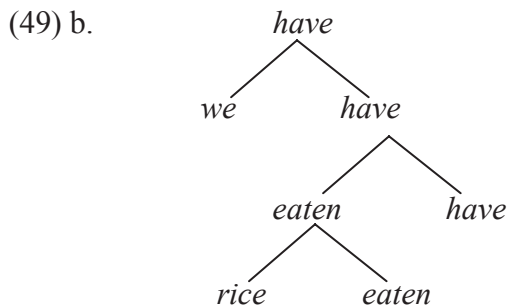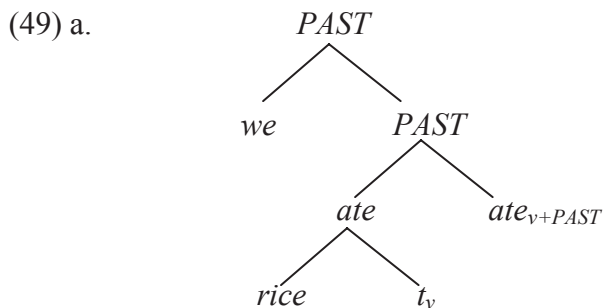
The postorder traversal in the BPS representations (49a, b) yields the following sequences, respectively.

(50) a. {*we*, *ate* + *PAST*, (*t*ᵥ,) *rice*, ~~*ate*~~,   b. {*we*, *have*, *eaten*, *rice*, *eaten*, **have**,
         PAST, *PAST*}   *have*}

And their respective reduction as below:

(51) a. {*we*, *ate*, *rice*, *PAST*}   b. {*we*, *rice*, *eaten*, *have*}

The perfect sentence (48b) does not match with (51b), which is apparently a problem.

Nevertheless, the problem is only apparent, and I claim that the perfect auxiliary is *v* that raises to I as in the following:

(52)
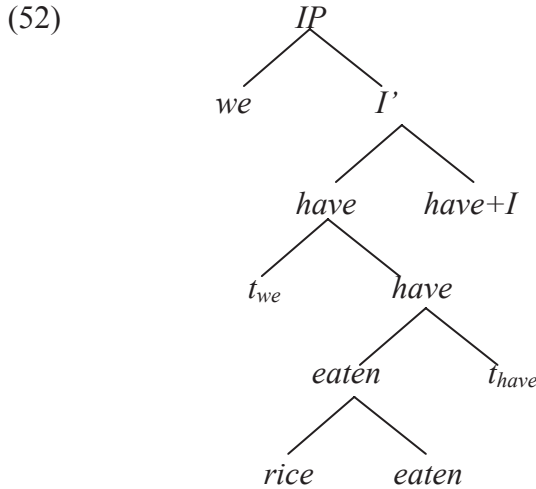
(53) {*we*, ***have + I***, (*t*$_{we}$,) (*t*$_{have}$,) *eaten*, ***rice***, ***eaten***, **have**, ~~**have**~~, **I′**, ***IP***}
(54) {*we*, ***have***, ***rice***, ***eaten***, ***IP***}

The postorder traversal in the BPS representation (52) yields the sequence (53) that in turn is spelled out as (54), which matches with (48b).

6. Concluding Remarks

As I have shown, the proposed BPS Traversals overcome the empirical as well as the theoretical problems left in Kural's (2005) tree traversal linearization. Dispensing with the directionality both in branching of structures and in traversal algorithms, word orders for *wh*-movement in VSO languages were accounted for, and typologically rare VOS, OVS, and OSV word orders are made underivable without movement. In addition, seemingly mixed orderings in Vata were also accounted for.

There is an important insight in Kural's proposal that it is the order of maximal projections that are reflected in word orders. In other words, the word orders are the "images" of the hierarchical relations of the maximal projections. It is almost the same intuition behind Kayne's (1994) LCA. The proposed

modifications also keeps the two important insights that Kayne's (1994) Theory of Antisymmetry implies; there should be no languages with specifiers on the right, and movement should be always to the left. The proposed BPS Traversals capture these insights without fixing the branching directions in phrase structures.

The proposed BPS Traversal linearization is a viable alternative to the LCA-based linearization. After all, the structures permitted by LCA are not as restricted as widely assumed. Guimarães (2008) has recently revealed that the following types of structures are allowed in principle: i) n-ary branching ($n > 2$); ii) heads adjoined to non-heads; iii) non-heads adjoined to heads; iv) multiple specifiers; and v) multiple adjunction to heads. Moreover, the Complement-Head-Specifier order is not logically excluded but by some obscure stipulation about the timing in linearization.

Although the proposed BPS Traversal linearization awaits further development with wider empirical coverage, I believe that the graph-theoretical traversal is a promising approach, in that its formal aspects are well-studied and it is domain independent. It is readily available to the Human Language Faculty by natural laws as the Third Factor in the sense of Chomsky (2005).

References

Chomsky, Noam. 1970. Remarks on nominalization. In *Reading in English Transformational Grammar*, eds. Roderick A. Jacobs and Peter S. Rosenbaum. Waltham: Ginn and Comapny.

Chomsky, Noam. 1994. *Bare Phrase Structure: MIT Occasional Papers in Linguistics* 5. Cambrdige, MA: MITWPL.

Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36:1–22.

Gaertner, Hans-Martin. 1995. Has bare phrase structure theory superseded X-bar Theory? *Forschungsschwerpunkt Allgemeine Sprachwissenschaft Papers in Linguistics* 4:22–35.

Guimarães, Maximiliano. 2008. A note on the strong generative capacity of standard Antisymmetry Theory. *Snippets* 18:5–7.

Jackendoff, Ray S. 1977. *X̄ Syntax: A Study of Phrase Structure*. Cambridge, MA: MIT Press.

Kayne, Richard S. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.

Koopman, Hilda. 1983. *The Syntax of Verbs: From Verb Movement Rules in the Kru Languages to Universal Grammar*. Dordrecht: Foris Publications.

Kural, Murat. 2005. Tree traversal and word order. *Linguistic Inquiry* 36:367–387.

McCloskey, James.1996. On the scope of verb movement in Irish. *Natural Language & Linguistic Theory* 14:49–104.

Yasui, Miyoko. 2003 A graph-theoretical reanalysis of bare phrase structure theory and its implications on parametric variation. To appear in *Proceedings of the International Conference on Linguistics and Phonetics 2002*. Prague: Charles University Press.

Gemination and Syllabic Repairs in Sanskrit and Middle Indic

Andrea Calabrese
University of Connecticut, Storrs

Abstract

This paper deals with consonantal gemination in Sanskrit and Middle Indic. It proposes that this gemination process resulted from a process called here root spreading. By spreading the root node of a segment, this process creates configurations that satisfy constraints against complex onsets or codas with the consequence that these syllabic configurations are eliminated. A detailed analysis of the development of consonantal clusters in Sanskrit and Middle Indic is provided.

1. Introduction

Saussure (1889) observed that there was a lack of contrast between the Sanskrit forms in (1a) as discussed in the passage in (1b)*:*

(1) a. *śaara-trayam* 'tree arrows' *śarattrayam (<śarad-trayam)* 'three autumns'
   b. Saussure (1889, 426-7):
   "Devant liquide, nasale ou semi-consonne, le catégories de la consonne double et de la consonne simple sont absolument confondues en sanskrit. Etant donnés les composés *śara-trayam* (trois flèches) et *śarad-trayam* (trois automnes), nous croyons devoir en Europe observer la differérence étymologique dans l'orthographe, écrire l'un *śaratrayam* et le second *śarattrayam*. Si nous consultons la tradition indigène, nous apprenons qu'il faut écrire:
   a.   D'après   nombre   de   manuscripts   dans   les   deux   cas *śaratrayam..* Aucune occlusive n'est marquée double devant [r, w].
   b.   D'après   certains   *Prātiśākhyas:*   dans   les   deux   cas *śarattrayam.* Aucune occlusive n'est marquée simple devant [r, w]."

Andrea Calabrese

     c.   D'après P`āṇini` (8, 4, 47); cf. avec critique 48 et 50-52): dans les deux cas *śaratrayam* ou dans le deux cas *śarattrayam*. Emploi à volonté de la lettre double ou simple devant [r, w].

Cette dernière doctrine, pour etre fidèlement rapportée, doit plutot se formuler comme suit: toute occlusive est supposée simple devant [r, w], mais on peut toujours la redoubler"

Saussure argued that the fundamental reason for the lack of contrast in the Sanskrit forms in (1)a) is that the stop in clusters stop+/sonorant was always geminated; this gemination was sometimes not represented orthographically in Sanskrit. The topic of this paper is this process of gemination. My hypothesis is that this gemination process results from a process removing complex onsets and codas. The basic assumption behind this hypothesis is that skeletal positions represent the interface between syllable structure and the melodic segments. A given melodic segment is assigned a given syllabic status through its association with a skeletal position. By changing its association relations with the skeletal positions we can change its syllabic status. Complex onsets are governed by constraints targeting the relations between classes of melodic segments and syllabic positions such as the one in (2a). We can then remove complex onsets by changing their sub-skeletal representation by an operation I will call root spreading as in (2b). (2b) eliminated complex onsets in Sanskrit (see (3)). As we will also see, the mirror image application of (2b) eliminated codas (see (4)):

(2) a. No complex onsets   b. Root spreading:

(3)

p   u   t   r   a   *p   u   t   r   a [puttra]* 'son'

(4)

$$\sigma \quad \sigma \qquad \sigma \qquad \sigma$$
$$R \quad R \qquad R \qquad R$$
$$N \quad N \qquad N \qquad N$$
$$X \; X \; X \; X \;\rightarrow\; X \; X \qquad X \; X$$
$$a \quad r \quad t^h \quad a \qquad a \quad r \qquad t^h \quad a$$

$[artt^ha]$ 'purpose'
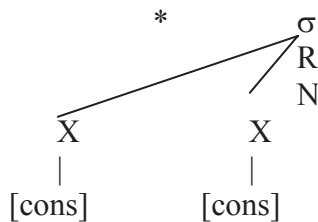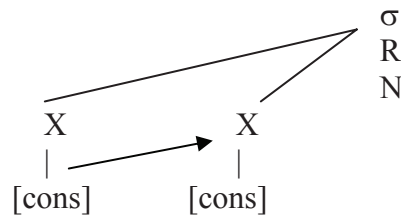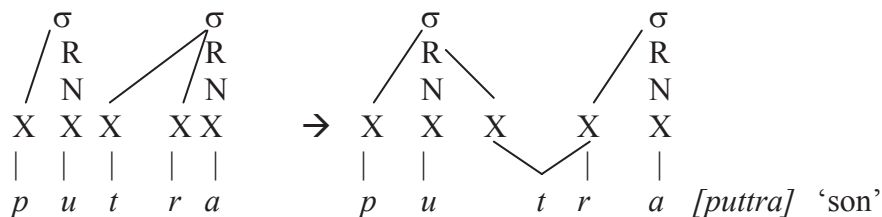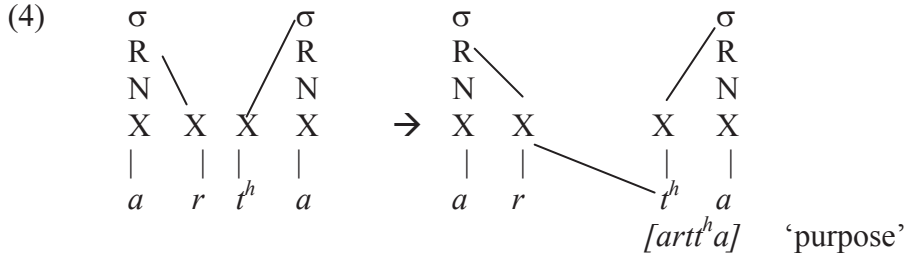
2. Sankrit and Middle Indic Gemination

In Calabrese (2009), I provide evidence showing that at least two different passes of syllabification must be postulated for Sanskrit. In an initial pass, complex onsets and simple and complex codas are allowed. In a later pass, in contrast, both complex onsets and simple and complex codas are disallowed. The main focus of this paper is the resyllabification process that brought about the changes characterizing the second pass of syllabification.

2.1 Graphic Doubling of consonants in Sanskrit Manuscripts

In many Sanskrit manuscripts, the first consonant of clusters of rising sonority is written as double as shown in (5) (see Vaux 1992, Whitney 1868, Wackernagel 1896:112) and especially Varma (1929, chapter II and V), for a detailed discussion of the facts).

(5)   *dádhy átra*   -->   *daddhy átra*   'sour milk + here '

  mádhv átra   -->   *máddhv átra*   'delicious + here'

  agní   -->   *aggní*   'fire'

  putráḥ   -->   *puttráḥ*   'son, child'

  satyáḥ   -->   *sattyáḥ*   'true'

Whitney (1868) provides the following account of the conditions under which graphic doubling occurs in Sanskrit:

(6) The first consonant of a cluster is doubled, but when the first member of the cluster belongs to the set *{r l v h}*, it is the second member which is doubled; in addition, the phonemes /r/ and /h/, geminates, and members of homorganic stop clusters do not geminate.

Calabrese (2009) shows that the first consonant of intervocalic clusters of rising sonority has a peculiar property in Sanskrit: it has a double identity: it seems to be the coda of the preceding syllable, thus making it heavy, but at the same time, it behaves as if it is in the onset of the following syllable. Observe that it is

Andrea Calabrese

precisely this consonant that is doubled in the manuscripts. Varma (1929) argues that this graphic doubling represents actual phonetic gemination: wherever a consonant is graphically double in the Sanskrit manuscripts, later stages of Sanskrit, and in particular Middle Indic, display an actual geminate in its place (see also Jacobi 1881: 609). We can thus assume that the Sanskrit word *putra* was phonologically represented as in (7):

(7)
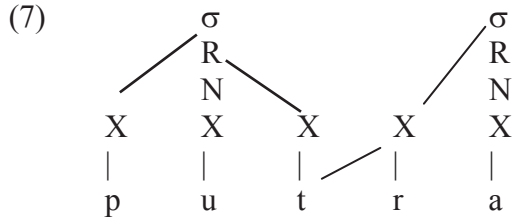
One of these middle Indic languages displaying cluster gemination is Pāli (Hankamer and Aissen 1974).

(8) a. Sanskrit: *kalya*　　　　Pāli:　*kalla*　　'ready'
　　　　　　catvaras　　　　　　*cattaro*　'four'
　　　　　　aśva　　　　　　　　*assa*　'horse'
　　　　　　kurvanti (v=[w])　　*kubbanti*　'they make'
　　　　　　kilbiṣa　　　　　　*kibbisa*　'guilty'
　　　　　　kalmāsa　　　　　　*kammāsa*　'freckled'

Notice, however, that clusters in word initial position are simplified by eliminating the second onset consonant (see later for discussion):

(9)　　Sanskrit:　　　　　　　　　Pāli:

　　　　kvaṭhati　'boils'　　　→　　*kaḍhai*
　　　　trasati　'tremble'　　→　　*tasati*
　　　　grāma　'village'　　→　　*gāma*
　　　　śruta　'heard'　　→　　*suta*
　　　　dhvani　'sound'　　→　　*dhani*
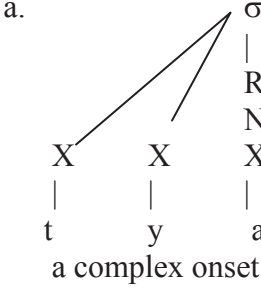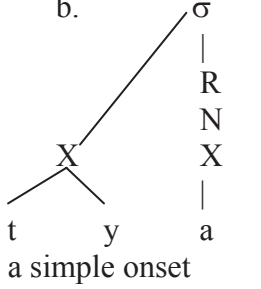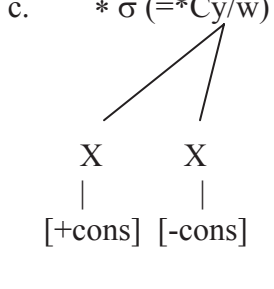
3. Explanation of Gemination in Sanskrit and Middle Indic

I propose that the main purpose of the gemination process we see in Sanskrit and Middle Indic is the elimination of complex onsets and codas.

Consider the status of skeletal positions and their relation to segmental roots. The skeletal positions must be considered as the interface level between the melodic component and syllable structure. A skeletal position represents the syllabic segment, the structural unit that is relevant for syllable structure. A skeletal position is to be distinguished from the root whose function is to represent the temporal overlap or simultaneity among the features it dominates. The root represents the melodic segment, a bundle of simultaneously articulated features, namely, the phoneme. Such a distinction between syllabic segments and melodic segments is needed to describe sounds such as the labio-coronal /ps/ of Margi (Ladefoged 1968), the velarized coronal affricate /tx/ of Xũ (Snyman 1970) or the double fricatives ꜰs, fʃ, ꞵʃ of SePedi (Ladefoged and Maddieson 1996). The languages where they are found have simple CV syllables and therefore do not allow consonantal clusters. Hence these sounds involve a single syllabic unit but two melodic elements. For example consider Margi /ps/ (see Clements 1992) (See Calabrese 2009) for more discussion and evidence in support of this type of representations):

(10)
```
                              s
                              R
                              N
            X                 X
         ╱     ╲
      Root     Root
       |        |
       p        s
```
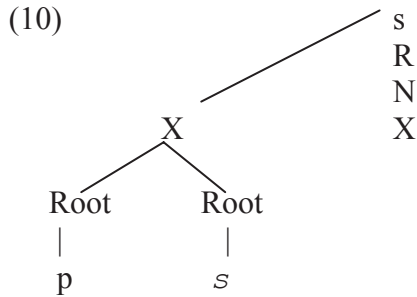
The crucial assumption in the analysis of gemination in Sanskrit is that a given melodic segment is assigned a given syllabic status through its association with a skeletal position. By changing its association relations to the skeletal positions we can change its syllabic status. Calabrese (2005) proposes that the operation of root spreading in (12) must be recognized as a strategy to repair syllable structure which can be used as an alternative to nucleus-insertion (epenthesis) and deletion. In particular, root spreading (2b) can repair violations of the constraint in (11c) and convert (11a) into (11b).

Andrea Calabrese

(11)   a.

$$\sigma$$
$$|$$
$$R$$
$$N$$
X   X   X
| | |
t   y   a

a complex onset

b.

$$\sigma$$
$$|$$
$$R$$
$$N$$
X   X
| |
t   y   a

a simple onset
(=the affricate in (10)

c.   * σ (=*Cy/w)

X   X
| |
[+cons] [-cons]
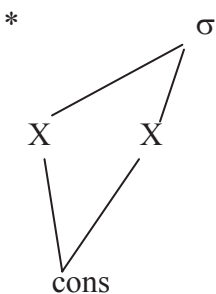
Consider the structure in (11a). If the root node of the first onset consonant is incorporated under the skeletal position dominating the glide as in (11b), we are no longer dealing with an onset cluster. To have an onset cluster one needs two or more skeletal positions, each exhaustively associated with a melodic segment. Thus, by the root spreading operation in (2b), the onset cluster is removed from the syllabic interface, and can no longer be targeted by the constraint in (11c).

As we can see in (12), the application of root spreading in (12) creates an onset geminate:

(12) Application of Root spreading to repair a complex onset:

s
R
N
a. X   X   X   X   -->   b. X   X   X   X
    | | | | | | | |
    V   C   y   V   V   C   y   V

s
R
N

s
R
N
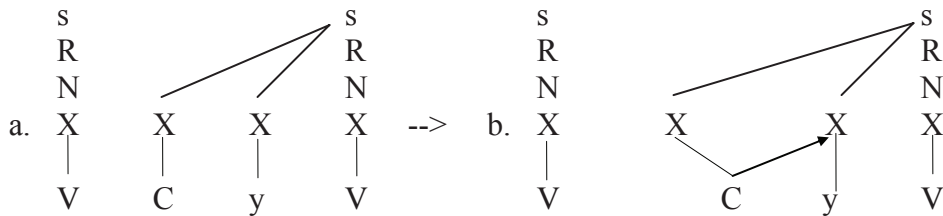
Onset geminates are disallowed by the constraint in (13). This onset geminate is repaired by delinking its first member from the onset position as in (14):

(13)   *   σ

X   X

cons

(14)

```
        s                     s         s                   s
        R                     R         R                   R
        N                     N         N                   N
  a.    X     X    X    X  --> b. X   X           X          X
        |     |    |    |         |                |          |
        V     C    y    V         V     C   y      V
```
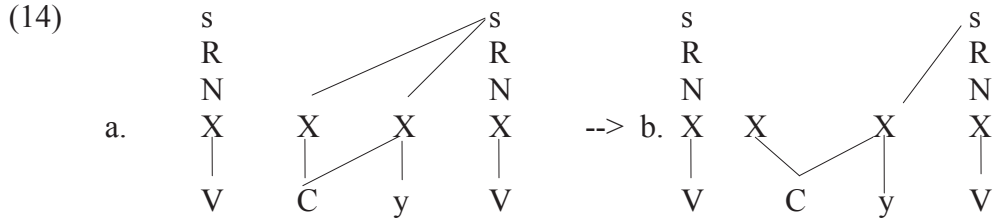
The unsyllabified consonant that results is incorporated into the preceding syllable thus resulting into a legitimate geminate.

(15)

```
        s                     s         s                   s
        R                     R         R                   R
        N                     N         N                   N
  a.    X     X    X    X  --> b. X   X           X          X
        |     |    |    |         |                |          |
        V     C    y    V         V         C  y   V
```
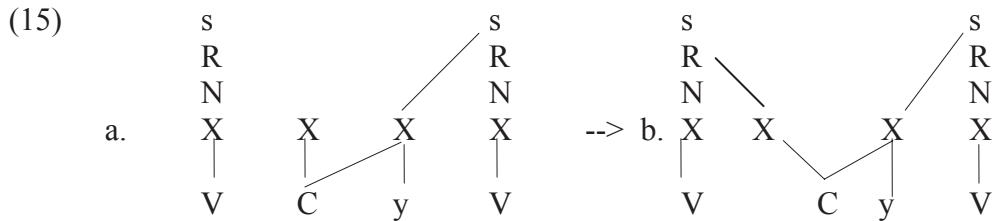
Summarizing what has been proposed above, we can repair an onset cluster by spreading the root of the first onset consonant onto the skeletal position of the following one. The surface effect of this repair is that of geminating the first consonant (see (16)):

(16)

```
  σ          σ       σ          σ          σ                σ
   R          R       R          R          R                R
   N          N       N          N          N                N
 a.X    X X   X --> b. X   X X   X --> c.   X  X       X      X
   |    | |   |        |   | | |   |        |   \     /|      |
   V    C R   V        V   C R   V          V      C R        V
```
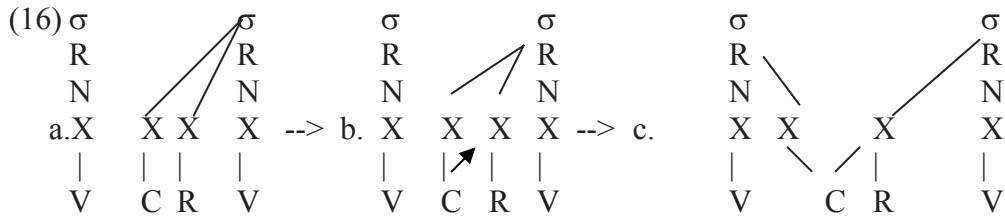
4. Analysis of Gemination in Sanskrit and Middle Indic

4.1 Elimination of complex onsets

Let us consider Sanskrit and Middle Indic again. Examples in (17) are from Pischel (1981:225, 233), Masica (1981:174-15).

(17)

a. Skt. Input		b. Skt. Surface	c. Middle Indi
aśakya	'impossible'	*aśakya*	*asakka*
supyatē	'sleeps'	*suppyatē*	*suppati*
ramya	'enoyable'	*rammya*	*ramma*
Sarasvati	'n. of goddess'	*Sarassvati*	*Sarassadi*
cakra	'wheel'	*cakkra*	*cakka*

aśru	'tear'	*aśśru*	*assu*
agnih	'fire'	*aggnih*	*aggi*
yugma	'pair'	*yuggma*	*jugga*

In Calabrese (2009), I proposed that in initial syllabification, complex onsets and codas are allowed in Sanskrit. In the preceding section, I argued that these complex onsets and codas are eliminated in a second pass of syllabification through the operation of root spreading as in (18).

(18)
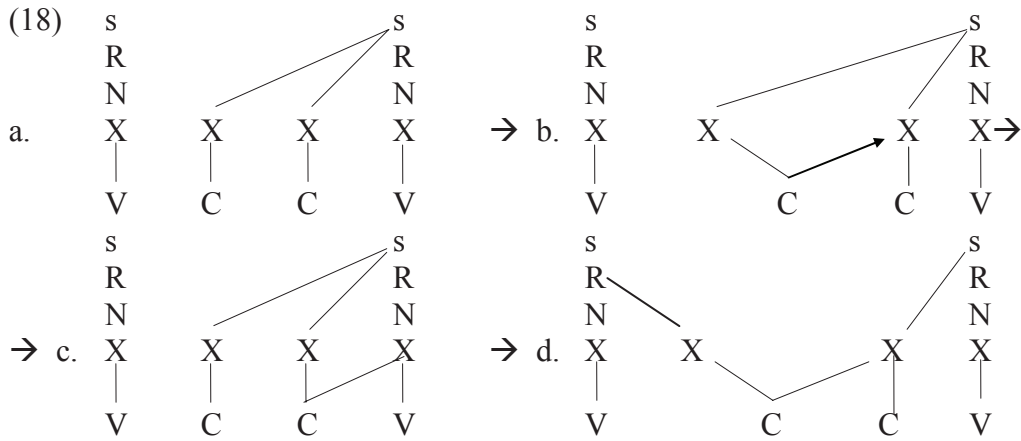
The Middle Indic outcomes can be accounted for by assuming a process delinking branching roots in (20). The Uniformity Applicability Condition (UAC) (Schein and Steriade 1986) (as reformulated in Calabrese (1999) in (19) allows the application of (20) only to singly-linked branching roots.

(19) Given a node *n* and a set *S* consisting of all nodes linked to *n,* and given a rule or   a constraint *T*, if *T* refers to *n* and any member of *S,* it must refer to all members of *S* to be active.

(20)

(21)

```
    s                              s   s                            s
    R \                          / R   R \                        / R
    N  \                        /  N   N  \                      /  N
    X       X          X      / X --> X       X          X      X
    |        \        /       /|              \          |     /
    V          C₁   C₂       V                 C₁
```
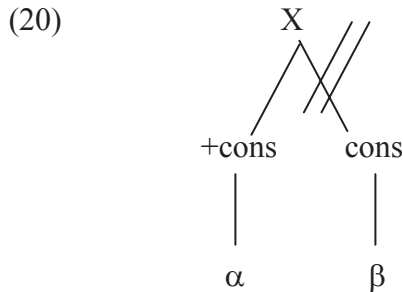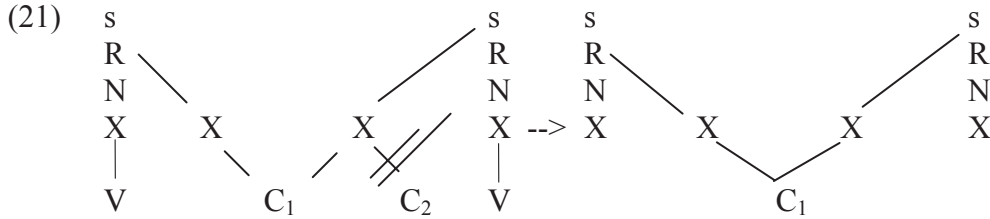
The analysis just proposed can be extended to word-initial position. No gemination in Sanskrit or in Middle Indic is observed in this case, as shown in (*e.g.*, Skt. *grāma* 'village' Middle Indic*:* *gāma*). I propose that in this case, the onset geminate resulting from the application of root spreading (2b) is eliminated by skeletal deletion (see (22b)):

(22) a.

```
            s                 b.              s               c.             s
          / |R                            / |R                            / R
         /  |N                           /  |N                           /  N
    X   X   X …  →            X   X   X …  →                X         X …
    |   |   |                 |   |   |                     / \       |
    g   r   a                 g   r   a                    g   r      a
```
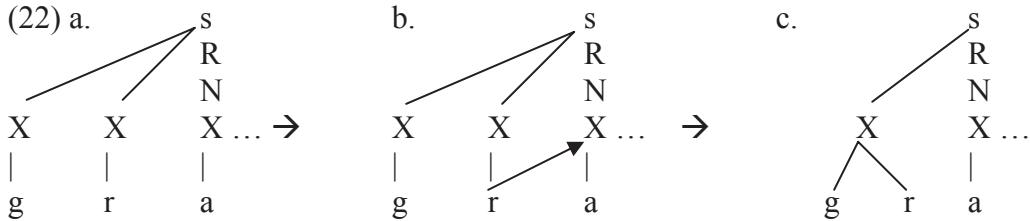
In the affricate-like structure in (22c), Middle Indic eliminates the most sonorous components as in (23):

(23) a.

```
            s                 b.                       s
          / R                                        / R
            N                                          N
    X   X …  →                            X       X …
    / \  |                                |       |
   g  r  a                                g       a
```
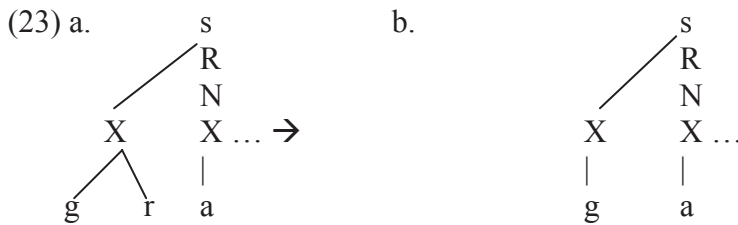
To account for the evolution of other consonantal clusters from Sanskrit to Middle Indic, some further rules are required, for example, aspiration. It is required to account for what happens in the clusters in (24) (Pischel (1981:258):[1]

(24) a. Skt. Input   b. Skt. Surface   c. Middle Indic
        kṣata   'wounded'   *kṣata*   *khata*

---

[1] Two rules of palatalization are also needed to account for the treatment of clusters in cases such as Skt. *satya* 'truth' Middle Indic *sacca* or Skt. *akṣi* 'eye' Middle Indic*:* *acchi* (when *k < ś*), Skt. *Apsaras* 'nymph' Middle Indic: *acchara*. They cannot be discussed here. (see Calabrese (2009))

aksi   'eye'   *akksi*   *akkhi*
bhiksu   'mendicant' *bhikksu*   *bhikksu*
śiksita   'learned'   *śikksita*   *sikkhida*

(25) Aspiration rule:

    [+cons]   [+cons]   //bidirectional
                             |
                     Laryngeal
                       |
             [+spread glottis]
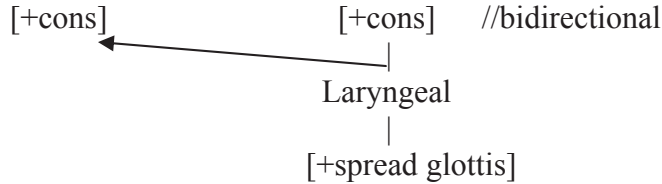
Consider a word such as *aksi*. (2b) generates (26b). This is the Sanskrit situation. In Middle Indic, rule (25) applied as in (27).

(26)

(27)
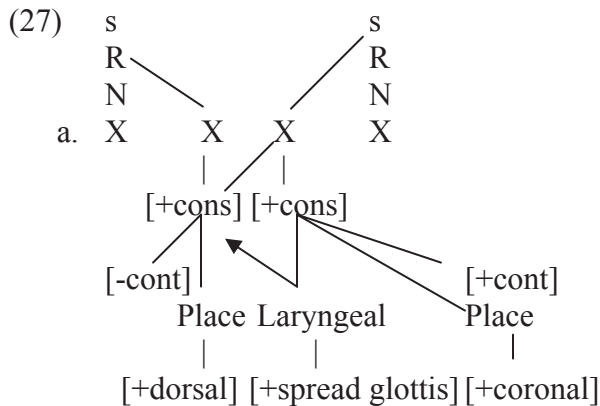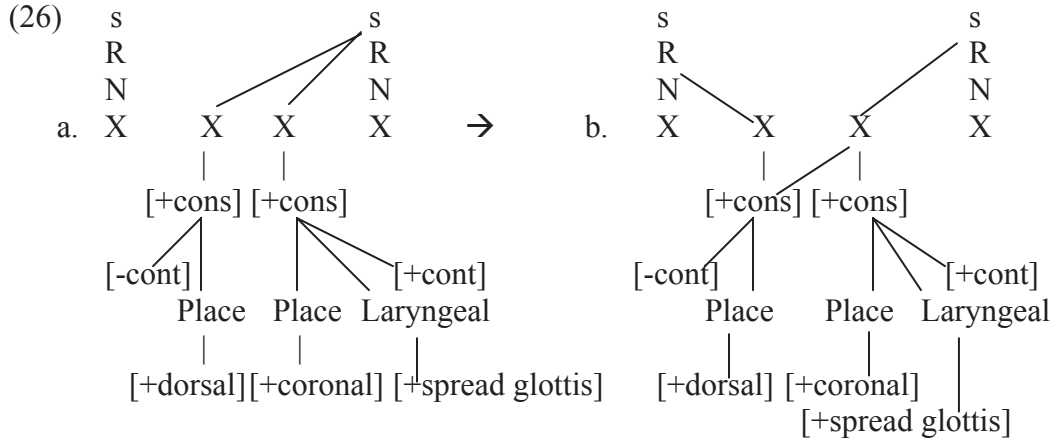
Subsequent application of (20) generates the geminated aspirated stop in (28):

(28)
```
       s                    s
       R                    R
       N                    N
   a.  X     X      X       X
             |
           [+cons]
         /   |
   [-cont]   |
         Place Laryngeal
           |      |
      [+dorsal] [+spread glottis]
```
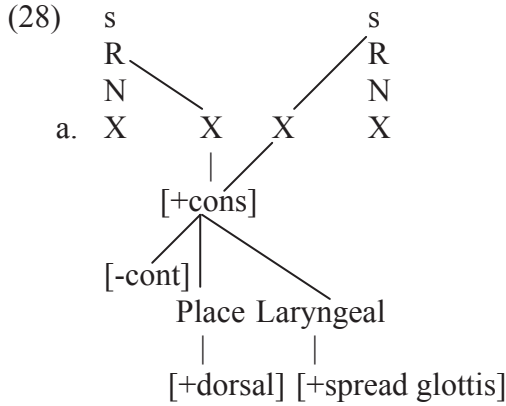
4.2   Elimination of Codas

Gemination is also found in clusters such as that in (29a) where we have a simple coda containing a liquid followed by simple onset. It is the simple onset that is geminated in this case. (See sample cases in (30) (Pischel 1981:233, Masica 1991:176)).

(29)
```
      s      s                  s        s
      R      R                  R        R
      N      N                  N        N
  a.  X  X  X  X           b.   X   X   X   X           {arttʰa}
      |  |  |  |                |   |       |
      a  r  tʰ a     -->        a   r   tʰ  a
```
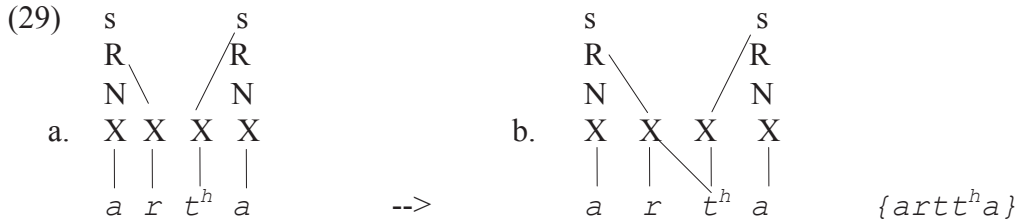
(30) a. Skt. Input   b. Skt. Surface   c. Middle Indic

Ardha	'half'	*arddha*	*addha*
mārga	'road'	*mārgga*	*magga*
artha	'purpose'	*arttha*	*attha*
arpita	'entrusted'	*arppita*	*appita*
alpa	'small'	*alppa*	*appa*

As discussed above, root spreading (2b) leads to the elimination of complex onsets in Middle Indic. I propose that its mirror image application also leads to the elimination of true codas where a true coda involves a violation of the constraint in (31) where we have an independent and unlinked root node.

Andrea Calabrese

(31)   R   (NO CODAS)
        N
                    X
                    |
                 [cons]

The only codas that are allowed in Middle Indic after the application of (2b) (mirror image) are the first member of a geminate as shown in (32). [2]

(32)   s   s
        R   R
        N   N
        X   X   X   X
                           [cons]

The constraint in (31) does not hold for (32) since the root in (32) is also part of the onset. Therefore, the coda in (32) does not violate (32) and is not a true coda. Therefore, the process in (2b) (mirror image application) eliminates true codas. [3]

    Consider the case of a cluster with [r] followed by a stop. The coda consonant violates (31). Root spreading (2b) (mirror image) repairs this violation as in (33). In fact, after the application of root spreading, the coda skeletal position in (33) is not affected by the constraint in (31) according to the UAC because it also dominates a root that is dominated by the onset of the following syllable. The Middle Indic outcome is obtained by the application of (20) as in (34):

(33)   s   s
        R   R
        N   N
        X   X   X   X
                             | |
                          +cons   [+cons]
                          +son
              [-nasal] | | [-cont]
                          Place   Place
                             |
                        [+coronal]

---

[2] Codas in Middle Indic can also contain nasals homorganic with the following onset. They behave like geminates, as discussed later.

[3] Observe that the final nasals, *i.e.*, the /m̥/ of `kattum̥` inf. kar 'to make', are anusvara which according to Calabrese (2009) are part of the syllabic Nucleus and therefore not true codas.

(34)

```
s                                              s
R                                              R
N                                              N
X                    X              X          X
                     |              |
                   +cons         [+cons]
                   +son                  [-cont]
          [-nasal]  |
                  Place          Place
                    |
               [+coronal]
```

Consider now clusters containing a coronal fricative followed by a stop. In sequences such as these there is gemination of the stop (Varma 1929:75, Pischel 1981:238, Masica 1991:177):

(35) *a.* Skt. Input   *b.* Skt. Surface   Middle Indic

hasta	'the hand'	*hastta*	*hattha*
vastu	'thing'	*vasttu*	*vatthu*
aṣṭā	'eight'	*aṣṭtā*	*attha*
puṣpa	'flower'	*puṣppa*	*puppha*
avaskanda	'assault'	*avaskkanda*	*avakkhand*

The coda consonant in the examples in (35) is disallowed by (31). Root spreading (2b) (in its mirror image application) removes the configuration disallowed by (31) as discussed above (see (36)).

(36)

```
s                                              s
R                                              R
N                                              N
X                    X              X          X
                     |              |
                  [+cons]        [+cons]
       [+cont]   Lar  Place    Place  [-cont]
                  |     |
          [+spread gl] [+coronal]
```
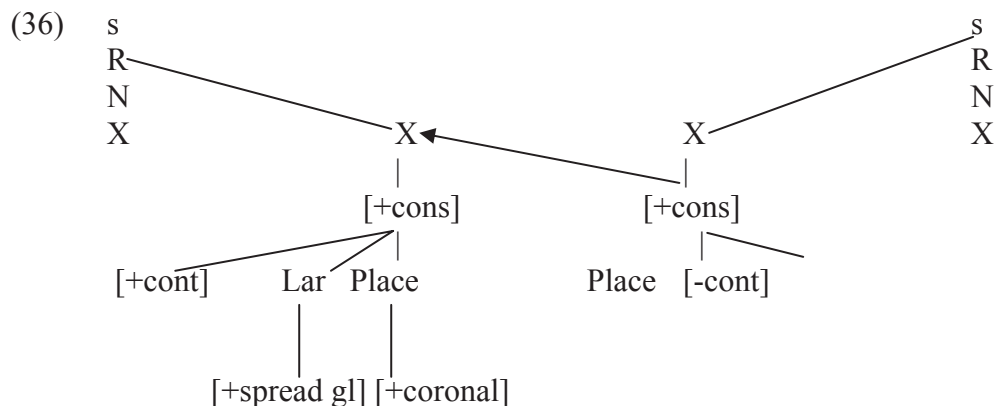
This accounts for the gemination we see in surface Sanskrit in these cases. The Middle Indic outcomes are derived by the application of (25) followed by the delinking operation in (20) thus generating (37b):

(37) a.
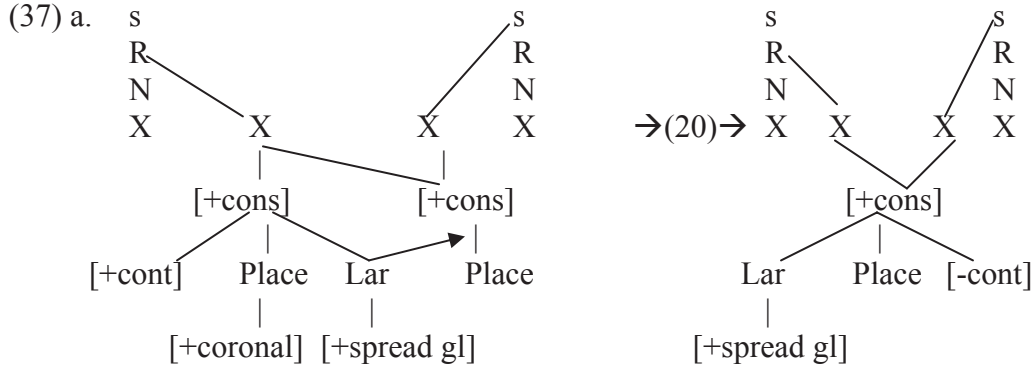
Root spreading in (2b) (in its mirror image application) immediately explains the gemination we find in stop clusters:

(38) a. Skt. Input   b. Skt. Surface   c. Middle Indic

	a. Skt. Input		b. Skt. Surface	c. Middle Indic
Bhakta	'meal, food'		*bhaktta*	*bhatta*
Dugdha	'milk'		*dugddha*	*duddha*
ṣaṭka	'set of six'		*ṣaṭkka*	*chakka*
utpāta	'sudden portent'		*utppāta*	*uppāta*
Sapta	'seven'		*suptta*	*sutta*
Labdha	'taken'		*labddha*	*laddha*

The input configuration for these clusters is shown in (39a). The stop in coda position is disallowed by (31). Root spreading applies to repair this configuration and geminates the second stops as in (39b). This is the Sanskrit surface situation. The Middle Indic outcomes are derived by applying (20), as in (39c):

(39) a. Sanskrit: b. Surface Skt.   c. Middle Indic
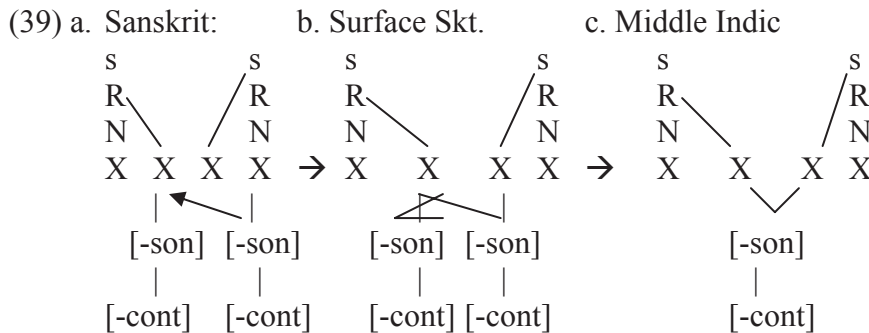
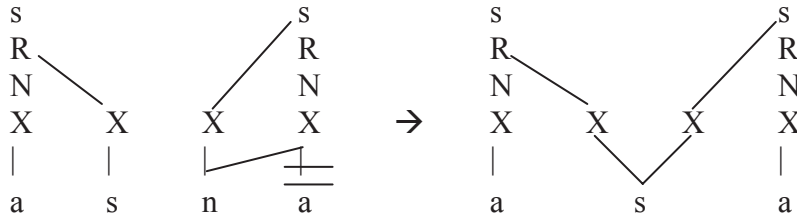We can consider now the sequence of a fricative plus a nasal. In initial syllabification, they behave as onsets (see Calabrese 2009). Here application of (2b) should give us gemination of /s/ (see (40) and (41) for the Middle Indic outcome with loss of the nasal due to (20).

(40)   Sanskrit Gemination
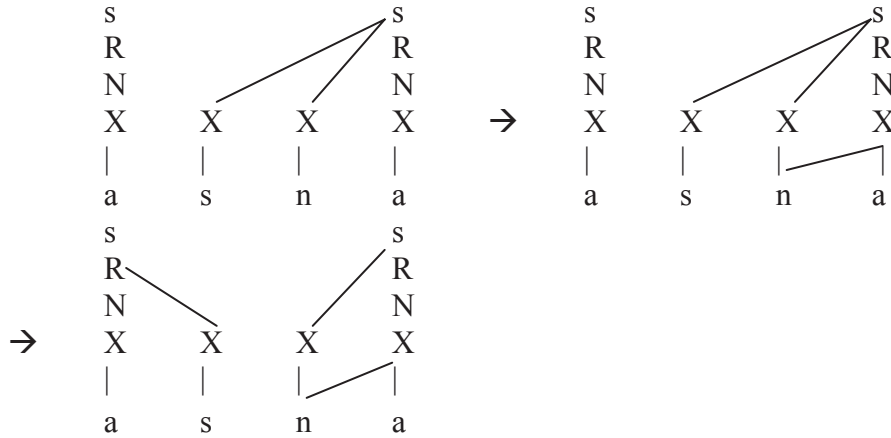
(41)   Middle Indic outcome

There are a few of such outcomes (see (42)). However, the most common development of this cluster in Middle Indic is a geminated aspirated nasal (see footnote below about the digraphs nh/mh)[4])

(42)   Skt. Input   Skt. Surface   Middle Indic
         raśmi   'rope'   rassmi   rassi
         etasmin   'this-Lsg.'   etassmin   edassim

(43)   *praśna*   'question'   *praśnna*   *paṇha*

---

[4] The digraph nasal+h represents an aspirated nasal (Masica 1981:178). Its geminate status is shown by the fact that it triggers shortening of the preceding vowel as all other geminates (Masica 1981:183): Skt. *grísma* 'summer heat' MI *gimha*; Skt. *ślēṣman* 'mucus, phlegm' MI silemha/silimha). This shortening is due to Geiger's Law according to which syllable rimes in MI cannot exceed two moras (see Calabrese 2009 for more discussion)

Andrea Calabrese

Kṛṣṇa	'dark blue'	*Kṛṣṇna*	*Kaṇha*
gṝiṣma	'summer heat'	*gṝiṣmma*	*gimha*
aśman	'stone'	*aśmman*	*amha*

Observe now that although onset clusters fricative+nasal need to be reconstructed for common Indo-European, they were eliminated in Indo-European languages such as Greek and Latin. They can be considered quite instable and marked onset clusters. We should expect a tendency to eliminate them before other clusters.
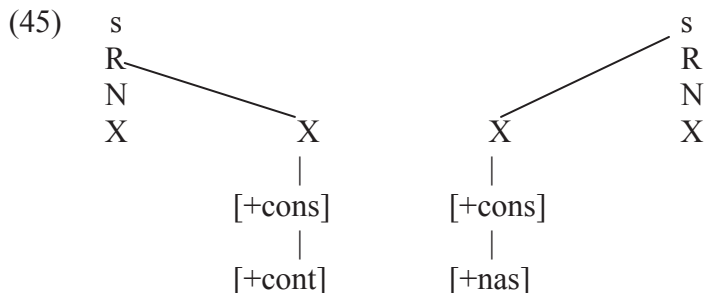
To account for the Middle Indic development we see in (42), I then propose that at a certain point of the history of Sanskrit, before the resyllabification process that lead to gemination occurred, there was a change in what was allowed in initial syllabification. In particular, onset clusters of a fricative followed by a nasal became disallowed, *i.e.*, the constraint in (44) became active).

(44)
$$*\qquad\qquad\text{s}\quad\text{(No s+nasal onsets)}$$

```
                 *                    s   (No s+nasal onsets)

           X               X
           |               |
        [+cons]         [+cons]
           |               |
        [+cont]         [+nas]
```
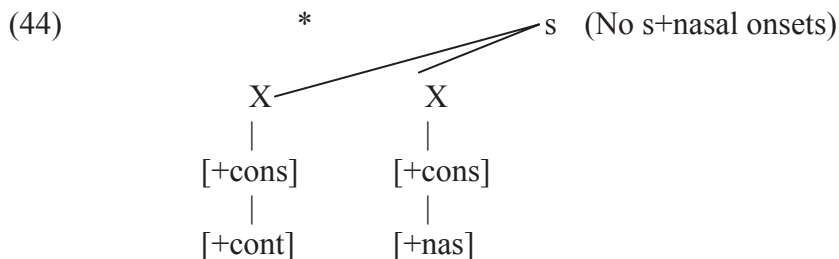
Therefore sequences of this type were syllabified as in (45) in initial syllabification

(45)
```
     s                              s
     R                              R
     N                              N
     X          X          X          X
                |          |
             [+cons]    [+cons]
                |          |
             [+cont]    [+nas]
```

In the second pass of syllabification, these configurations were removed by root spreading as in (46):

(46)
```
s                              s
R────────────                  R
N            ╲                 N
X             X      X         X
              │      │
          [+cons]  [+cons]
              │      │
          [+cont]  [+nas]
```
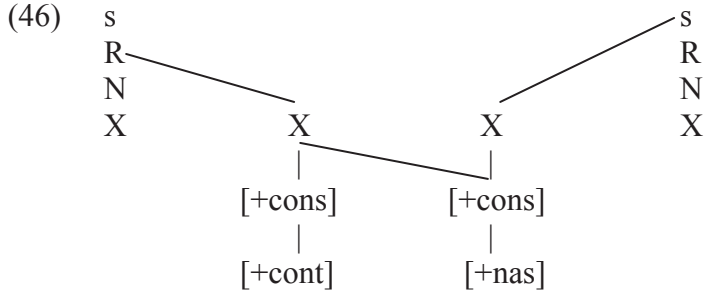
The Middle Indic outcomes are accounted for by applying the rule of aspiration in (25) as shown in (47). (20) then applies and (48) is generated. [5]

(47)
```
s                              s
R────────────                  R
N            ╲                 N
X             X      X         X
              │      │
          [+cons]  [+cons]
              │   ──────→│
     [+cont] Laryngeal      [-cont]
              │      │
              │   [+nasal]
     [+spread glottis]
```

(48)
```
s                              s
R────────────────────          R
N                    ╲         N
X                     X    X   X
                           │
                       [+cons]
                      ╱    │   ╲
              Laryngeal  [+nasal] [-cont]
                  │
          [+spread glottis]
```
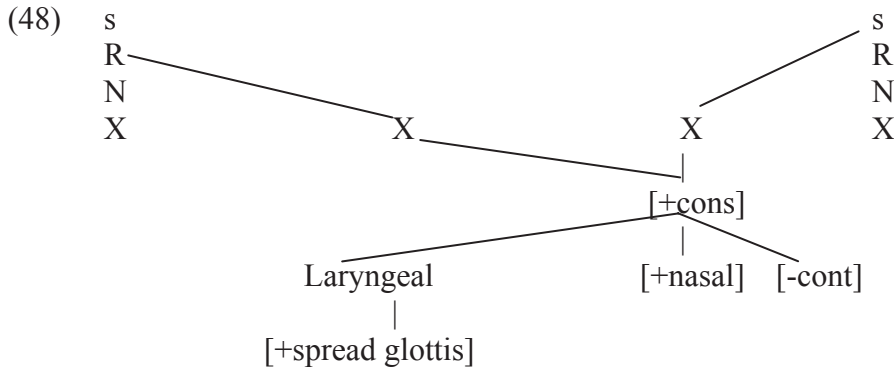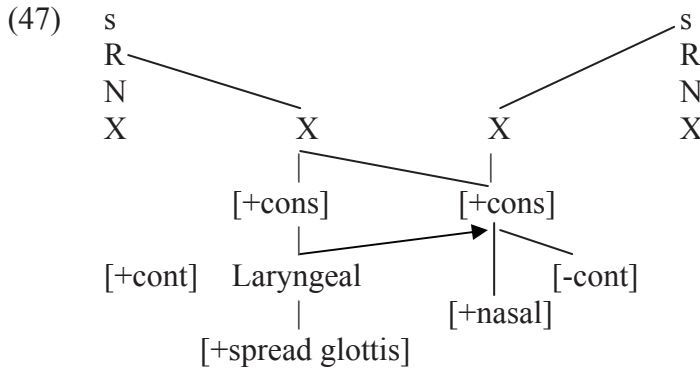
[5] Root spreading followed by skeletal deletion applies when these clusters occur in word-initial position: (*snāti* 'bathes' → *nhāi, snāru* 'sinew' → *nhāru, snūsā* 'son'swife' → *nhusā*). In the case of these clusters also epenthesis can be observed (*snigdha* 'sticky' → *siniddha / saniddha*). There are also cases with deaspiration: *snigdha* 'sticky' → *ṇiddha*)

Andrea Calabrese

No graphic doubling in Sanskript manuscripts and no gemination in Middle Indic is found in the case of homorganic nasal stop clusters:

(49) a. Skt. Input   b. Middle Indic

antara	'interior'	*antara*
andha	'blind'	*andha*
aŋkuśa	'elephant'	*aŋkusa*
lamba	'pendent'	*lamba*

Nasals followed by a homorganic stop have the structure in (50):

(50)
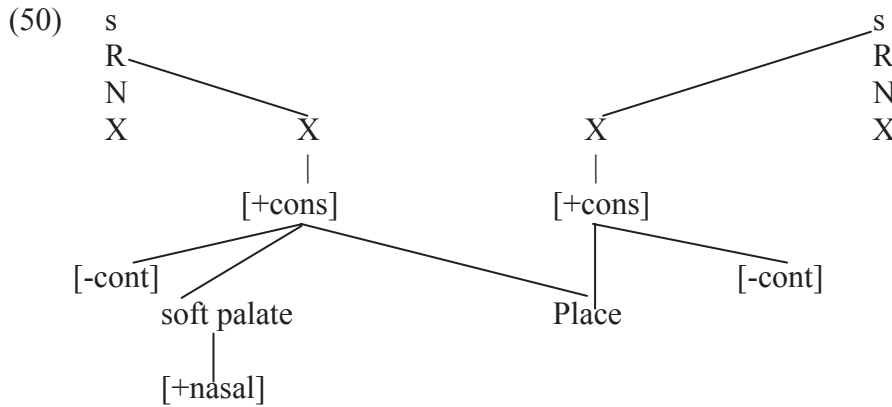
This structure is produced by a previous application of a process of nasal place assimilation given in (51).

(51)
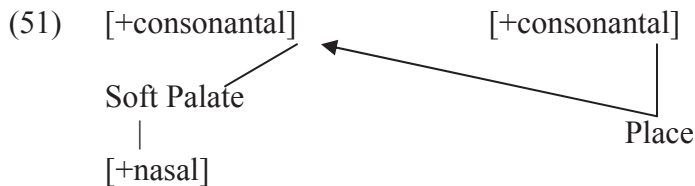
I propose that the structure in (50) is automatically changed into (52). In other words, I propose that a homorganic nasal+ stop sequence is automatically changed into a prenasalized geminate stop, and that phonetically there is no distinction between them. Observe that I assume that prenasalized stops have the same affricate-like structure of the complex segments as in (10) with two root nodes linked to a single skeletal position (see Clements 1992, Calabrese 2005: chapt.4, 2009):

(52)
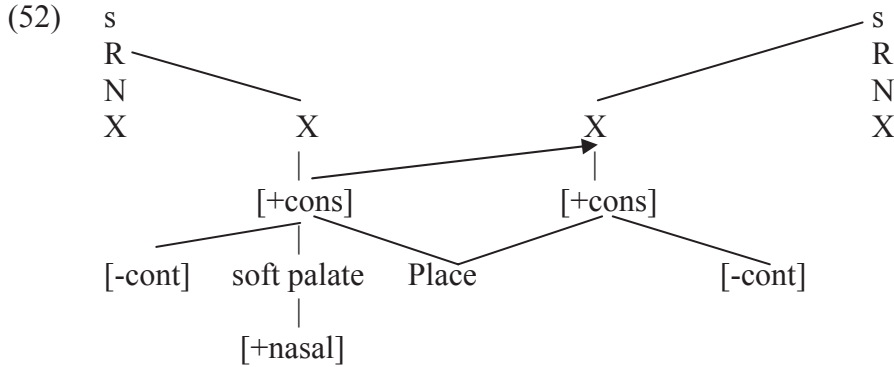
The coda skeletal position is therefore licensed by the onset root according to the UAC. No root spreading is then needed in this case. Therefore, there is no gemination in this case in Sanskrit and subsequently these structures are preserved as such in Middle Indic.[6, 7]

5. Conclusion

In this paper, I have shown that if we assume that an operation like root spreading can repair disallowed syllabic configurations, we can readily account for the gemination processes we find in Sanskrit and Middle Indic. Root spreading must be recognized as another possible source for gemination, in addition to processes such as assimilation, lengthening under stress, reanalysis (see Blevins 2005)

References

Blevins, Juliette. 2004 *Evolutionary phonology*. Cambridge: Cambridge University Press, 2004.

Calabrese, Andrea. 2005. *Markedness and Economy in a Derivational Model of Phonology*. De Gruyter-Mouton,

Calabrese, Andrea. 2009. Cluster syllabification, Gemination and Syllabic repairs in Sanskrit. Ms. University of Connecticut

Clements, Nick G. 1992. Place of articulation in consonants and vowels: A unified theory.In *L'Architecture et la Géometrie des Repré sentations Phonologiques*, B.Laks and A.

---

[6] I assume that (20) does not apply in the prenasalized stop configuration with a single place node in (52).

[7] See Calabrese (2009) for an analysis of Sanskrit polisegmental clusters such as those in the following forms: Skt. *tīkṣna* 'sharp' MI. `tinha/ tikkha` (Pischel (1981:254) Skt. *lakṣmi* 'good fortune' MI. `lacchi.`

Andrea Calabrese

Rialland, (eds.), Paris (Also appeared in *Working Papers of the Cornell Phonetics Lab* 5 (1991)).

Hankamer, Jorge and Judith Aissen. 1974. The Sonority Hierarchy. *Papers from the Parasession on Natural Phonology*, A. Bruck, R. Fox, and M. La Galy, eds. Chicago: Chicago LinguisticSociety.

Ladefoged, Peter. 1968. *A Phonetic Study of West African Languages*. Cambridge: Cambridge University Press.

Ladefoged Peter and Ian Maddieson. 1996. *The Sounds of the World Languages*. Cambridge, MA: Blackwell.

Jacobi, H. 1881 "Zur Genesis der Präkritsprachen" *Zeitschrift fur Vergleighende Sprachforschung* 25.603-609.

Pischel, R. 1981. (transl. by Subhadra Jha) *A grammar of the Prākrit Languages* Dehli: Motilal Barnasidass.

Masica, Colin P. 1981. *The Indo-Aryan Languages* Cambridge: Cambridge University Press.

Saussure, Francois, de. 1889. "Sur un point de la phonétique des consonnes en indo-européen." in *Receuil de publications scientifique*, Geneva 1922: Sonor, 420-432.

Schein Barry and Donca Steriade. 1986. "On Geminates." *Linguistic Inquiry* 17. 691-744.

Snyman, Jan. 1970. *An Introduction to the !Xũ languages*, Balkema, Cape Town.

Varma, Siddheshwar. 1929. *Critical Study in the Phonetic Observations of the Indian Grammarians.* London: Royal Asiatic Society.

Vaux , Bert. 1992. "Gemination and Syllabic Integrity in Sanskrit: *The Journal of Indo-European Studies. 20,* 3-4.

Wackernagel, Jakob. 1896. *Altindische Grammatik , I.* Göttingen: Vandenhoek und Ruprecht.

Whitney, William .D. 1868 *The Taittirīya Pratiśākhya and Tribhāṣyatna.* New Haven: Yale University Press; reprinted Delhi: Motilaal Barnasidass.

On Children's *Need*s

Koji Sugisaki

Mie University, Japan

Abstract

This study evaluates the recent crosslinguistic proposal by Harves & Kayne (2008) and Harves (2008a,b) that the languages that have a close counterpart of the transitive verb *need* are limited to a proper subset of those languages that have a close counterpart of the transitive *have*. Our results from child English argue for the analysis by Harves & Kayne (2008) and Harves (2008a,b) in which the transitive *need* is derived through an incorporation to a silent counterpart of the transitive *have*. The findings of this study suggest that the time course of child language acquisition is a potentially rich source of evidence about the nature of variation permitted by human language.

1.   Introduction

Within the Principles-and-Parameters approach to UG (including the recent Minimalist Program), a theory of crosslinguistic variation is simultaneously a theory of the child's "hypothesis space" during language acquisition. The task for a child is to identify the correct grammar for the community's language from among the possibilities permitted by UG. In principle, then, we can gain insight into the nature of permitted variation by investigating how the child's grammar changes during the course of acquisition.

Recent studies of child language convincingly demonstrated that child language acquisition is indeed a valuable source of evidence concerning possible cross-linguistic variation, especially in the domain of syntax. For example, Snyder (2001) revealed that English-learning children acquire endocentric root compounds and transitive verb-particle constructions at around the same time, and proposed a syntactic parameter that correlates these two distinct properties. Sugisaki & Snyder (2005/2006) provided evidence from child English for Kayne's (1981) parametric proposal that the availability of preposition stranding is linked to the availability of the prepositional complementizer construction. In

addition, Sugisaki (2008) showed that the course of acquisition is consistent with the analysis (such as Hasegawa 2007) that creates an implicational relationship between the languages that permits the swiping construction (Merchant 2002) and the languages that allows preposition stranding.

In this study, I develop this line of research, and attempt to add another piece of evidence that child language acquisition constitutes an important testing ground for evaluating hypotheses about language variation. In contrast to the earlier studies, the present study focuses on an (apparent) cross-linguistic variation in the lexicon. More specifically, this study evaluates the recent proposal by Harves & Kayne (2008) and Harves (2008a,b) that, among Indo-European languages, the languages that have a close counterpart of the transitive verb *need* are limited to a proper subset of those languages that have a close counterpart of the transitive verb *have*. Our results from child English argue for the analysis by Harves & Kayne (2008) and Harves (2008a,b) in which the transitive *need* is derived through an incorporation to a silent counterpart of the transitive *have*.

2. *Have, Need, and Want*: Accounting for their Crosslinguistic Variation

It has been widely known that transitive verbs like *need* and *want* appear to take a simple DP argument which receives an intensional interpretation. The sentences involving these transitive verbs exhibit two hallmark characteristics of intensional contexts first discussed by Frege (1892): (i) Substitution of co-referring terms need not preserve truth, and (ii) non-denoting objects need not induce falsity. For example, substituting "Spiderman" in (1a) for "Peter Parker" in (1b) need not preserve truth, even though these two terms are extensionally equivalent. The sentence in (2) remains true even though "a unicorn" has no reference in the actual world. In contrast, regular transitive verbs do not exhibit these properties: Substitution of "Spiderman" for "Peter Parker" in (3) preserves truth, and the sentence in (4) is simply false.

(1) a. MJ needs/wants Spiderman.
    b. MJ needs/wants Peter Parker.

(2) Ken needs/wants a unicorn.

(3) a. MJ met Spiderman.
    b. MJ met Peter Parker.

(4) Ken owns a unicorn.

Intensional transitive verbs *need* and *want* share other properties as well (see Larson, den Dikken, and Ludlow 1997 and Schwarz 2007, among others). For example, both of these verbs exhibit ambiguities with time adverbial adjuncts, as illustrated in (5) (Schwarz 2007).

(5)   Matt needed/wanted some change before the conference.

    a.   Paraphrase 1:
        There was a time before the conference at which Matt needed/wanted some change.

    b.   Paraphrase 2:
        What Matt needed/wanted is to have some change before the conference.

However, a recent, detailed comparative survey of Indo-European languages by Harves & Kayne (2008) and Harves (2008a,b) revealed that, despite such similarities, these two transitive verbs significantly differ in their cross-linguistic distribution. They observe that the availability of a transitive verb corresponding to *need* is severely limited in that it is present only in *H(ave)-languages*, the languages that have a close counterpart of English *have*, which is an overt verb expressing ordinary possession such that the possessor has nominative Case and the possessee is a direct object (with accusative Case and no preposition). In other words, transitive *need* is not available in *B(e)-languages*, the languages that lack a close counterpart of *have* and express ordinary possession using some counterpart of *be*. For example, Czech is an H-language and has a transitive verb corresponding to *need*, while Russian is a B-language and lack a transitive *need*: As illustrated in (7b-d), Russian expression that corresponds to English *need* either requires a Dative-Nominative pattern or requires a preposition in front of its object.

(6)   Czech (Harves & Kayne 2008:5-7, Harves 2008b:8):

    a.   *Possession*
        Mají   nové   auto.
        have-3PL   new   car-ACC
        'They have a new car.

    b.   *'Need'*
        Tvoje   děti   tě   potřebují.
        your   children-NOM   you-ACC   need-3PL
        'Your children need you.'

c. *'Want'*

Petr	chce	nové	auto.
Petr-NOM	want-3SG	new	car-ACC

'Petr wants a new car.'

(7) Russian (Harves & Kayne 2008:3-4, Harves 2008a:215):

a. Possession

U	Ivana	budet	novaja	mašina.
at	Ivan-GEN	be-FUT.3SG	new	car-NOM.SG

'Ivan will have a new car.'

b. 'Need'

Mne	nužna	èta	kniga.
me-DAT	necessary-FEM	that	book-NOM.FEM

'I need that book.'

c. 'Need'

Rebenok	nuždaetsja	v	vašej	pomošči.
child-NOM	need-3SG	in	your	help-PREP

'The child needs your help.'

d. 'Need'

*Rebenok	nuždaet	vašu	pomošč.
child-NOM	need-3SG	your	help-ACC

e. 'Want'

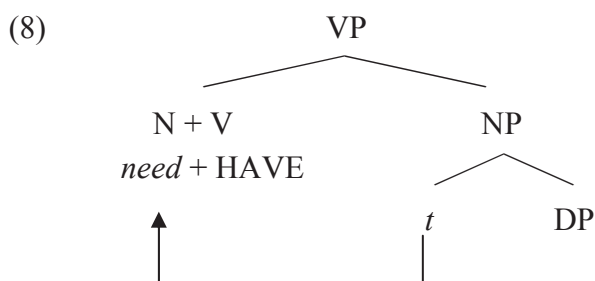| Ivan	xočet	novuju	mašinu.
Ivan-NOM	want-3SG	new	ca-ACC

'Ivan wants a new car.'

Harves (2008a,b) observes that, in sharp contrast to the transitive *need*, a transitive verb corresponding to English *want* can be present both in H-languages and in B-languages. For example, both Czech and Russian have a counterpart of English *want*, even though the former is an H-language and the latter is a B-language, as illustrated in (6c) and (7e), respectively.

Part of the comparative survey by Harves & Kayne (2008) and Harves (2008a,b) is summarized in Table 1. In order to account for the crosslinguistic generalization in Indo-European languages that a transitive verb corresponding to *need* is available only in H-languages, Harves & Kayne (2008) adopt a Hale & Keyser (1993) style incorporation-approach to verbal *need*, and propose a derivation in which nominal *need* raises and incorporates into unpronounced verbal HAVE. They argue that the incorporation of nominal *need* into silent

HAVE results in the appearance of a transitive *need* inheriting the accusative Case licensing properties of HAVE.

	Possession?	Transitive *need*?	Transitive *want*?
English	have	YES	YES
German	have	YES	YES
Spanish	have	YES	YES
Czech	have	YES	YES
Icelandic	have	YES	NO
Swedish	have	YES	NO
French	have	NO	YES
Italian	have	NO	YES
Bulgarian	have	NO	YES
Russian	be	NO	YES
Hindi	be	NO	YES
Irish	Be	NO	NO
Welsh	Be	NO	NO

Table 1: Crosslinguistic Survey of Indo-European Languages

(8)

```
                        VP
              _____
             |                |
           N + V             NP
         need + HAVE      _____
                         |        |
                         t        DP
```

By assuming that the presence of a silent verb in a language requires an overt counterpart, the lack of transitive *need* in B-languages follows straightforwardly from the derivation shown in (8): Since silent HAVE constitutes a necessary component in the derivation in (8), transitive *need* is available only in those languages that can expresses possession using transitive *have*. Put another way, B-languages lack transitive *need* precisely because they lack transitive *have* and by extension its silent counterpart.

Koji Sugisaki

As we can see in Table 1, there are H-languages that lack transitive *need*, which indicates that the presence of transitive verbal *have* is a necessary but not sufficient condition for licensing verbal *need* in a given language. For example, French has transitive *have* (and *want*) but still lacks transitive *need*, as illustrated in (9).

(9)   French (Harves & Kayne 2008:13, Harves 2008b:6-7):
   a.   Possession
         J'ai   une voiture.
         I-have-1SG   a   car
         'I have a car.'

   b.   'Need'
         J'ai   besoin   d'une   voiture.
         I-have-1SG   need   of-a   car
         'I need a car.'

   c.   'Want'
         Je   veux   une voiture.
         I   want-1-SG   a   car
         'I want a car.'

Harves (2008b) speculates that the restricting factor would be the availability of an incorporation operation: The H-languages that disallow transitive *need* would be those that lack incorporation altogether. Then, in order for a language to have verbal *need*, the language must permit both a silent counterpart of transitive *have* and an incorporation operation.

   To summarize, Harves & Kayne (2008) and Harves (2008a,b) observe that, even though transitive *want* can be present both in H-languages and in B-languages, transitive *need* is available only in a subset of H-languages. To account for this cross-linguistic distribution of verbal *need* in Indo-European languages, they proposed an analysis in which transitive *need* involves incorporation into the silent counterpart of transitive *have*.[1]

---

1.   In order to account for the intensionality effects observed with *need* and *want*, Harves (2008b) proposed that these transitives select for a hidden *v*P complement, building on the Small Clause analysis by Schwarz (2007). Crucially, under the analysis by Harves, languages differ as to which verb is embedded within this clausal complement: In H-languages, this structure involves silent HAVE or GET, while in B-languages, the structure involves silent BE. See Harves (2008b) for details.

3.   Prediction for the Acquisition of English

Even though Harves & Kayne (2008) and Harves (2008a,b) are quite careful in limiting their generalization to Indo-European languages, one may criticize their comparative survey by saying that the sample size is relatively small, by the standards of language typology. In addition, the diagnostics employed are quite superficial, and it would be necessary to become more precise about what counts as the counterpart of "transitive *need*" and of "transitive *have*".

In light of these potential problems, evidence from acquisition can be especially valuable: We can overcome these difficulties by deriving and testing the acquisitional prediction of their proposal. According Harves & Kayne (2008) and Harves (2008a,b), among Indo-European languages, natural-language grammars permitting transitive *need* are a proper subset of those permitting transitive *have*. In terms of acquisition, this generalization suggests that the knowledge of verbal *have* constitutes a proper subset of the knowledge required for deriving verbal *need*. Then, in the acquisition of a language like English that has both transitive *have* and *need*, when the presence of *have* is acquired earlier than other prerequisites for transitive *need* (such as the availability of an incorporation operation), children will acquire transitive *have* earlier than transitive *need*. On the other hand, when the presence of *have* is the last-acquired prerequisite for transitive *need*, children will acquire transitive *have* and *need* at around the same time. Hence, the proposal by Harves & Kayne (2008) and Harves (2008a,b) makes the following ordering prediction for the acquisition of English.

(10) Prediction for the Acquisition of English:
    In English, any given child will acquire transitive *have* prior to, or at around the same time as, but never significantly later than transitive *need*.

The comparative survey by Harves (2008a,b) suggests that, in contrast to verbal *need*, there should be no derivational link between transitive *have* and transitive *want*. Then, we can expect that children learning English will not exhibit any strict ordering effect between *have* and *want*. In the next section, we investigate the acquisition of transitive *have*, *need*, and *want* in English, with the goal of evaluating the prediction in (10).

4.   Transcript Analysis

In order to determine the validity of the prediction in (10), I analysed ten longitudinal corpora for English from the CHILDES database (MacWhinney

Koji Sugisaki

2000), which provided a total sample of more than 246,000 lines of child speech. The corpora analysed in this study are summarized in Table 2. For each child, I located the first clear uses of (i) transitive *have*, (ii) transitive *need*, and (iii) transitive *want*. To count as a clear use, these verbs were required to be followed by an overt DP object. The CLAN program Combo, together with a file of all the relevant forms of these predicates, was used to identify potentially relevant child utterances, which were then searched by hand and checked against the original transcripts to exclude imitations, repetitions, and formulaic routines.

Child	Collected by	Age Span	# Child Utterances
Adam	Brown (1973)	2;03 – 4;10	45,555
Anne	Theakston et al. (2001)	1;10 – 2;09	19,902
Aran	Theakston et al. (2001)	1;11 – 2;10	17,193
Becky	Theakston et al. (2001)	2;00 – 2;11	23,339
Eve	Brown (1973)	1;06 – 2;03	11,563
Naomi	Sachs (1973)	1;02 – 4;09	15,960
Nina	Suppes (1973)	1;11 – 3;03	31,505
Peter	Bloom (1970)	1;09 – 3;01	26,891
Sarah	Brown (1973)	2;03 – 5;01	37,012
Shem	Clark (1978)	2;02 – 3;02	17,507

Table 2: Corpora Analysed

Results are summarized in Table 3. All ten children produced transitive *have*, *need* and *want* by the end of their corpora. Following Stromswold (1996) and Snyder (2007), the age of acquisition was taken as the first clear use, followed soon after by repeated use. Mean age of acquisition for transitive *have* was 2;01 (years:months), with a range of 1;06 to 2;05. Mean age of acquisition for transitive *need* was 2;03, with a range of 1;09 to 2;08. Mean age of acquisition for transitive *want* was 2;00, with a range of 1;06 to 2;06.

To evaluate the statistical significance of observed age differences between acquisition of transitive *have* on one hand and acquisition of transitive *need* and *want* on the other, I counted the number of clear uses of the earlier construction before the first clear use of the later construction. Next I calculated the relative frequency of the two constructions in the child's own speech, starting with the transcript after the first use of the later construction, and continuing through the end of the corpus. A Binomial Test was then used to obtain the probability of sampling the observed number of tokens of the earlier construction simply by

chance, before the first clear use of the later construction. The null hypothesis for the test is that the second construction was grammatically available at least as early as the first construction, and had the same relative frequency observed in later transcripts (Stromswold 1996, Snyder 2007).

Child	Transitive *have*	Transitive *need*	Transitive *want*
Adam	2;04.03	2;06.17	2;06.03
Anne	1;11.08	2;02.12	1;10.17
Aran	2;02.25	2;08.12	1;11.12
Becky	2;02.30	2;03.27	2;02.15
Eve	1;06	1;09	1;06
Naomi	1;10.18	1;10.18	1;09.26
Nina	2;00.24	2;02.12	2;01.06
Peter	2;00	2;01	2;00
Sarah	2;05.30	2;08.02	2;03.07
Shem	2;03.16	2;03.02	2;02.16
Mean	2;01	2;03	2;00

Table 3: Ages of Acquisition

The results of the statistical analysis are summarized in Table 4. Four children acquired verbal *have* significantly earlier than verbal *need*, and the remaining six children acquired these two predicates at approximately the same age (no significant difference, $p > .10$). Crucially, no child in this study acquired transitive *need* significantly earlier than transitive *have*, as predicted in (10). In sharp contrast, in the case of transitive *have* and *want*, two children acquired transitive *have* significantly earlier than transitive *want*, six children acquired transitive *have* significantly later than transitive *want*, and the remaining two children acquired these two predicates at approximately the same age. Hence, the results of my transcript analysis indicate that even though the order of acquisition between transitive *have* and transitive *want* can vary from child to child, the order of acquisition between transitive *have* and transitive *need* is quite strict: Children never acquire transitive *need* significantly earlier than transitive *have*.

A possible alternative explanation for the observed order of acquisition in English would be to say that the order merely reflects the frequency of these predicates in the child-directed speech. In order to determine whether such input-based account can be maintained, I analyzed the same ten corpora and counted the number of sentences involving transitive *have*, *want*, and *need* in the

Koji Sugisaki

mother's utterances, up to the point when the child begins using all of these transitive verbs.

| Child | Transitive *have* vs. Transitive *need* | | Transitive *have* vs. Transitive *want* | |
	Early acquired	*p =*	*Early acquired*	*p =*
Adam	have	$p < .05$	Have	$p < .05$
Anne	have	$p < .01$	Want	$p < .001$
Aran	have	$p < .001$	Want	$p < .001$
Becky	at around the same time	$p > .05$	Want	$p < .001$
Eve	at around the same time	$p > .10$	Have	$p < .05$
Naomi	at around the same time	-----	Want	$p < .001$
Nina	have	$p < .001$	at around the same time	$p > .05$
Peter	at around the same time	$p > .10$	at around the same time	-----
Sarah	at around the same time	$p > .10$	Want	$p < .01$
Shem	at around the same time	$p > .10$	Want	$p < .001$

Table 4: Results of the Statistical Analysis

Child	Files analyzed	Transitive *have*	Transitive *need*	Transitive *want*
Adam	01-08	54	12	17
Anne	01a-14a	244	33	129
Aran	01a-27a	539	85	162
Becky	01a-12a	132	17	93
Eve	01-08	196	13	59
Naomi	01-15	35	1	39
Nina	01-14	159	2	21
Peter	01-07	12	0	15
Sarah	001-024	58	3	47
Shem	01-04	7	1	7

Table 5: Frequency of *have, need,* and *want* in the child-directed speech

The results are summarized in Table 5. In the adult speech, transitive have tended to be most frequent and transitive need tended to be least frequent. Since there were children who acquired verbal have significantly later than verbal want, and

also children who acquired verbal have and need at around the same time, the frequency in the adult input cannot be regarded as the crucial factor that explains the order of acquisition of these predicates.

5. Conclusion

The results of my transcript analysis greatly strengthen the generalization by Harves & Kayne (2008) and Harves (2008a,b) that, at least among Indo-European languages, natural-language grammars permitting transitive *need* are a proper subset of those permitting transitive *have*. The findings lend acquisitional support to their analysis that transitive *need* is derived through incorporation to a silent counterpart of transitive *have*, and consequently, to a decompositional approach to predicates (Hale & Keyser 1993). A broader implication of this study is that the time course of language acquisition provides an important testing ground to evaluate hypotheses about crosslinguistic variation, which in turn suggests that child language is potentially a valuable source of evidence concerning the nature of variation permitted by human language (Snyder 2001, 2007, Sugisaki 2003).

References

Bloom, Lois. 1970. *Language Development: Form and Function in Emerging Grammars*. Cambridge, Massachusetts: MIT Press.

Brown, Roger. 1973. *A First Language: The Early Stages*. Cambridge, Massachusetts: Harvard University Press.

Clark, Eve. 1978. Strategies for Communicating. *Child Development* 49:953-959.

Frege, Gottlob. 1892. Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik* 100:25-50.

Hale, Kenneth, and Samuel Jay Keyser. 1993. On Argument Structure and the Lexical Expression of Syntactic Relations. In *The view from Building 20: Essays in linguistic in honor of Sylvain Bromberger*, eds. Kenneth Hale and Samuel Jay Keyser, 53-109. Cambridge, Massachusetts: MIT Press.

Harves, Stephanie. 2008a. Intensional Transitives and Silent HAVE: Distinguishing between *Want* and *Need*. In *Proceedings 27th West Coast Conference on Formal Linguistics*, eds. Natasha Abner and Jason Bishop, 211-219. Somerville, Masachusetts: Cascadilla Proceedings Project.

Harves, Stephanie. 2008b. The Syntax of Intensional *Want* and *Need*: What's HAVE Got to Do with It? Ms., Pomona College.

Harves, Stephanie, and Richard Kayne. 2008. Having Need and Needing Have in Indo-European. Ms., Pomona College and New York University.

Hasegawa, Hiroshi. 2007. Swiping Involves Preposition Stranding, Not Pied-Piping. A paper presented at GLOW 30, University of Tromsø, Norway.

Kayne, Richard. 1981. On Certain Differences between French and English. *Linguistic*

Inquiry 12:349-371.

Larson, Richard, Marcel den Dikken, and Peter Ludlow. 1997. Intensional Transitive Verbs and Abstract Clausal Complementation. Ms.

MacWhinney, Brian. 2000. *The CHILDES Project: Tools for Analyzing Talk*. Mahwah, NJ: Lawrence Erlbaum.

Merchant, Jason. 2002. Swiping in Germanic. In *Studies in Comparative Germanic Syntax*, eds. C. Jan-Wouter Zwart and Werner Abraham, 295-321. Amsterdam: John Benjamins.

Sachs, Jacqueline. 1983. Talking about Here and Then: The Emergence of Displaced Reference in Parent-Child Discourse. In *Children's Language, Volume 4*, ed. K. E. Nelson. Hillsdale, New Jersey: Lawrence Erlbaum.

Schwarz, Florian. 2007. On *needing* Propositions and *looking for* Properties. In *Proceedings from Semantics and Linguistic Theory XVI*, eds. Masayuki Gibson and Jonathan Howell, 259-276. Ithaca, New York: CLC Publications, Cornell University.

Snyder, William. 2001. On the Nature of Syntactic Variation: Evidence from Complex Predicates and Complex Word-Formation. *Language* 77:324-342.

Snyder, William. 2007. *Child Language: The Parametric Approach*. New York: Oxford University Press.

Stromswold, Karin. 1996. Analyzing Children's Spontaneous Speech. In *Methods for Assessing Children's Syntax*, eds. Dana McDaniel, Cecile McKee and Helen Smith Cairns, 23-53. Cambridge, Massachusetts: MIT Press.

Sugisaki, Koji. 2003. *Innate Constraints on Language Variation: Evidence from child language*. Doctoral dissertation, University of Connecticut, Storrs.

Sugisaki, Koji. 2008. Swiping in Child English. In *Proceedings of the 27th West Coast Conference on Formal Linguistics Poster Session*, ed. Kevin Ryan, 71-77: UCLA Working Papers in Linguistics.

Sugisaki, Koji, and William Snyder. 2005/2006. The Parameter of Preposition Stranding: A View from Child English. *Language Acquisition* 13:349-361.

Suppes, Patrick. 1973. The Semantics of Children's Language. *American Psychologist* 88:103-114.

Theakston, Anna. L., Elena V. M. Lieven, Julian M. Pine, & Caroline F. Rowland. 2001. The Role of Performance Limitations in the Acquisition of Verb-Argument Structure: An Alternative Account. *Journal of Child Language* 28: 127-152.

Variation in English Free Choice Items

Veneeta Dayal

Rutgers University, New Brunswick

Abstract

This paper has two related goals. One is provide an account of variation among Free Choice Items (FCI) in English, FC *any, -ever* free relatives and *some N or other*. The other is to give a uniform account for two problems in the analysis of FC *any,* subtrigging and the partitive restriction. All FCI are claimed to have a multi-dimensional meaning, sharing their truth conditional contribution with standard one-dimensional quantifiers but encoding on top of that, a requirement of indeterminacy. FC *any* requires strong indeterminacy: there cannot be a single set of individuals that comprises the extension of the relevant property in every accessible world. It is predicted to be unacceptable in precisely those cases where the truth conditional meaning contradicts this requirement of fluctuation. Subtrigging and the partitive restriction is shown to follow from this. In contrast to the indeterminacy requirement of *any*, the indeterminacy requirement of *-ever* and *some N or other* is weak, merely implying lack of knowledge on the part of the speaker about the identity of individuals or lack of relevance of their identity.

1. Overview

There has been significant progress in our understanding of Free Choice *any* and of *-ever* free relatives in recent years.[1] Very often, however, the investigation of these two Free Choice Items (FCI) in English has been conducted in parallel. One of the aims of this paper is to bring the two strands together in a way that can account for commonalities and differences between them and, perhaps, a third FCI *some N or other* whose semantics has not so far been studied. The second is to account for two problems in the analysis of FC *any* that have proved recalcitrant to theoretical explanation.

I claim here that all FCI have a multi-dimensional meaning, sharing their truth conditional contribution with standard one-dimensional quantifiers like

---

[1] Many, but not all, of the major recent contributions on these topics are in the references. Thanks to audiences at the University of Amsterdam, the University of Massachusetts-Amherst, Georgetown University, University of Pennsylvania and Asian GLOW 7. Special thanks to Gennaro Chierchia for probing questions and comments on an earlier version.

every, plain free relatives and *some*, but encoding on top of that, a requirement of indeterminacy.

I characterize the indeterminacy of FC *any* as a grammatical constraint against there being a single set of individuals that have the relevant property in every accessible world. It is predicted to be unacceptable in precisely those cases where the truth conditional meaning contradicts this requirement of fluctuation. This allows for a uniform account of subtrigging and the partitive restriction. I also take note, in accounting for the behavior of *any* in negative statements, of supplemental *any* which has distinct properties from regular FC *any*. I attempt a compositional account of the multi-dimensional analysis which, if correct, not only will give theoretical bite to FC *any's* inherent modal feel, it might also deliver its known proclivity for wide scope.

In contrast to the indeterminacy requirement of *any*, the indeterminacy requirement of *-ever* and *some N or other* is weak, merely implying lack of knowledge on the part of the speaker about the identity of individuals or lack of relevance of their identity. Such indeterminacy is compatible with there being no single set of individuals across all relevant worlds (as is the case with *any*) or with the existence of such a set.

I should state at the outset that I take this to be an analysis in progress. There remain various loose ends to be tied up and extensions to be explored. Some of these I list in the conclusion. I also discuss in this connection, albeit very briefly, an influential recent approach to free choice effects.

2. FC *Any*: the Core Generalizations

Several explanations have been given for the distribution of English FCI *any* but there does not exist at present any single account that captures the full range of facts. The problem can be demonstrated by the following paradigm:

(1) Any owl / *any of these owls hunts mice.

(2) a. Bill read any book *(he found) / *(that was on his reading list).
    b. *Bill read any of these books/any of the books he found.

(3) a. Bill may/*must read any of these books.
    b. Bill may read any book (he finds)/(on his reading list).
    c. Bill must read any book *(he finds)/ *(on his reading list).

(1) illustrates the canonical generic use of FC *any*. Note that the partitive is unacceptable here.[2] That is, FC *any* can participate in generic statements but not in habitual statements. Data like this are at the heart of the intuition that the domain of *any* extends beyond a contextually defined set, accounting for the continued importance of the notion of widening, proposed by Kadmon and Landman (1993).

(2) illustrates the subtrigging effect, discussed first in Legrand (1975). FC *any* is not good in episodic statements without an appropriate relative clause or other phrasal modifier. In certain contexts, the modification may be covert (see Dayal 1998). But partitive *any* remains unacceptable, even with an appropriate modifier present.[3]

The importance of partitive *any* emerges again in modal contexts, illustrated in (3). Possibility modals are a hospitable environment for *any* but not necessity modals. As in the episodic case, necessity modals take non-partitive FC *any* with subtrigging but they cannot accept partitive *any*. The generalization, then, is as given in (4):

(4)	Generic	◇-Modal	□-Modal	Episodic
Unmodified Any	√	√	*	*
Subtrigged Any	√	√	√	√
Partitive Any	*	√	*	*

Dayal (1998) noted subtrigging and partitivity as problematic for Kadmon and Landman, highlighting in different ways the limitations of widening as an explanatory notion. Modification of the common noun in subtrigging should not affect the strengthening requirements for widened domains while the inner definite in partitives should block widening altogether.

The alternative account proposed there can be illustrated by considering necessity and FC *any*. To explain subtrigging, I treated *any* as a universal whose domain of quantification includes all possible individuals. (3c), for example, is

---

[2] There may be a reading in which reference is to sub-kinds of owls (Lisa Selkirk p.c.).

[3] The judgments assume a contextually anchored reading for the inner DP of the partitive, clearer in the case of demonstratives than definites. The judgments for definites only hold to the extent that a definite has a rigid interpretation. Thanks to Graham Katz and Paul Portner for discussion of this point. The ameliorating effect of subtrigging correspondingly is dependent on the modifier not anchoring the reference contextually.

unacceptable without a modifier because it is an impossible command -- there are books in worlds in which Bill does not exist. The choice of a very wide domain clashes with the imperative which has to be restricted to worlds within Bill's reach. The modifier introduces a spatio-temporal bound on the domain of quantification, making it possible for the command to be fulfilled. The distribution of *any* is also subject to an overarching requirement of vagueness, which says that the set of individuals who end up with the property denoted by the verb phrase should not be contextually salient. Note that (3a) does not encode an impossible command, the books being firmly grounded in the here and now by the partitive. But the combination of universal quantification over worlds and universal quantification over a fixed set of books violates vagueness: there can be no doubt about the set of books that will be read by Bill if he fulfils the command.

This two-pronged proposal has been criticized by various scholars. The explanation for the subtrigging effect has been thought to be too weak for the kind of ungrammaticality that is perceived, the explanation for the partitive restriction for being non-compositional. And, of course, there is unease about the need for two unconnected parts to the explanation for the licensing of *any*. At the same time, it would be fair to say that no account of FC *any* has been developed since that work that presents a truly satisfactory alternative to both problems. This, then, is another attempt at tackling what has remained elusive in our understanding of FC *any*.[4]

3. FC *Any*: the Proposal in a Nutshell

I would like to propose that *any* is a universal quantifier, like *every* (and *each*). Unlike *every*, however, *any* also has a conventional implicature of *fluctuation* (F-implicature), which states that no single set of individuals is such that it constitutes in every accessible world the set of individuals with the two relevant properties in that world. FC *any* is ruled out in statements whose truth conditional meaning contradicts F-implicature: $D_{FC}(P)(Q) = D(P)(Q) + q_{implicature}$ so that: $*D_{FC}(P)(Q)$ iff $D(P)(Q) \Rightarrow \neg q$:[5]

---

[4] The account of subtrigging and the partitive restriction presented here is a partial reversal of the claims in Dayal (1998) and can be seen as a development of ideas present in Dayal (1995a).

[5] Note that in the case of partitive *any*, the property P will be built out of the phrase *of these/the N*, with *of* of type $<e<s<e,t>>>$. It will denote a function from worlds w' to the set of entities that are individual parts in w' of the plural entity denoted by the definite inner DP at w-a:

(5) a. ⟦Any⟧ $=\lambda P \lambda Q \forall x [P(w)(x) \rightarrow Q(x)]$ *Universal Quantifier*

   b. $\neg \exists X \forall w'.w_a \leq w'. \lambda x [P(w')(x) \& Q(w')(x)] = X$ *F-implicature*

I take *every* and *any,* to be universal quantifiers that differ only in the presence or absence of this implicature: *any* implicates fluctuation, *every* does not.[6] The distribution of *any* is restricted but that of *every* is not, because *any* is incompatible with non-fluctuating contexts. The notion of fluctuation differs crucially from the notion of widening in making reference not only to the nominal property but also to the verbal property. This yields the distribution of FC *any*, as we will see below.
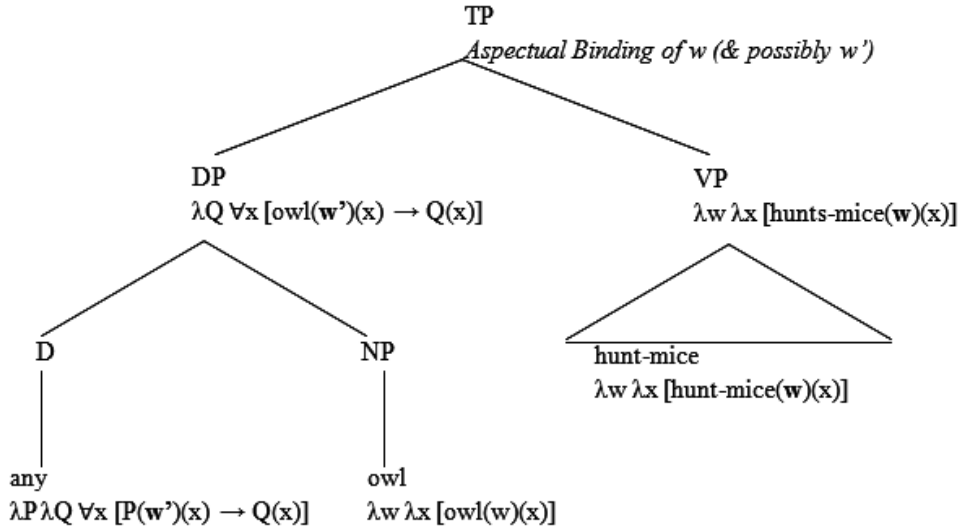
4. Deriving FC *Any*: Generic Statements

The quintessential case of FC *any* is a generic statement where the set of individuals denoted by the noun phrase varies with worlds. Recall that partitive *any* is not acceptable in such statements. Put another way, a partitive obligatorily involves a habitual reading and *any* is not acceptable in habitual statements. Understanding the difference between generic statements and habitual statements is therefore a good starting point for our account of FC *any*.

I make the standard assumption that the world/spatio-temporal variable on the VP is bound by tense and aspect in the extended projection of the verb, while the world/spatio-temporal variable on the DP has the option of remaining free. This gives us the schema in (6) and the possible derivations in (7a)-(7c) for the cases under discussion. I am assuming that genericity involves some form of universal quantification over worlds/situations:

$\lambda w' \lambda x [x \leq_{w'} iota(N(w_a)))]$. The extension of this function in any world, then, will be the contextually salient set N at w_a. Whether this set will be in the extension of N in the worlds of the modal base will depend on the modality involved (see also ft 8). I am indebted to Gennaro Chierchia for extensive discussion on this point, which has led hopefully to greater clarity in the articulation of the idea of fluctuation as applied to partitive *any*.

[6] Note that *any* and *each* are not in complementary distribution, cf: *Bill may read each/any of these books*. Clearly, *each* places restrictions only on the set denoted by the noun (the set N must be contextually anchored). *Any* is sensitive to the nominal as well as the verbal property.

Veneeta Dayal

(6)

```
                          TP
              Aspectual Binding of w (& possibly w')
                 /                          \
              DP                              VP
    λQ ∀x [owl(w')(x) → Q(x)]      λw λx [hunts-mice(w)(x)]
       /              \                     /\
      D               NP          _____/  _____
      |                |          hunt-mice
     any              owl         λw λx [hunt-mice(w)(x)]
λPλQ ∀x [P(w')(x) → Q(x)]   λw λx [owl(w)(x)]
```

(7) a. ∀w w$_a$ ≤$_{generic}$w ∀x [owl(**w**)(x) → hunts-mice(**w**)(x)] *any, every*

　b. ∀w w$_a$ ≤$_{generic}$w ∀x [owl(**w$_a$**)(x) → hunts-mice(**w**)(x) *every/each; *any*

　c. ∀w w$_a$ ≤$_{generic}$w ∀x [x ≤$_{w-a}$ ιy [owl(**w$_a$**)(y)] → hunts-mice(**w**)(x)] *partitive any*

　　The truth conditional contribution of an *any*-statement and an *every*-statement is the same. If w' in (6) is bound by verbal operators, we get a generic statement, with owls varying with worlds. If w' in (6) remains free, we get a habitual statement about the set of owls in a given world. A habitual statement contradicts the requirement of fluctuation that accompanies *any*. *Any*-statements are therefore necessarily generic while *every*-statements can be generic or habitual. *Each*-statements are only habitual:[7]

　　We see how the implicature of fluctuation can be used to derive the acceptability of *any* in generic statements and its unacceptability in habitual statements. The generic context, however, poses the least challenge for theories of *any* and there are other approaches that deal equally well with them. Let us

---

[7] The generic reading of universals has not featured prominently in the literature on genericity, but see Saeboe (2001) for relevant discussion. I should emphasize that I do not treat *any* as a generic indefinite, for reasons that I have discussed previously in Dayal (1998) and (2004).

turn now to the more complex cases and see how an explanation in terms of fluctuation fares in those contexts.

5. Deriving FC *Any*: Modal Statements

Let us start with the robust generalization that partitive *any* is only acceptable with possibility modals. As noted earlier, the fact that partitive *any* is sensitive to the modal in the verbal projection shows clearly that an explanation local to the nominal domain cannot work. Widening, rooted as it is in the nominal property, is inoperative in these cases. The notion of fluctuation, however, makes the right distinctions.

Consider (8a). It is easy verified that a wide scope universal over possibility allows for different books to be read in different worlds, satisfying fluctuation: $\{<w_1, \{a\}>, <w_2, \{a,b\}>, <w_3, \varnothing>, <w_4, \{b\}>\}$:[8]

(8) a. Bill may read any of these books.

b. $\forall x \ [x \leq_{w\text{-}a} \ \iota y \ [book'(w_a)(y)] \rightarrow \exists w'.w_a \leq w'. \ [read'(w')(x)(b)]]$

c. <u>book a</u>: $w_1 \rightarrow a, w_2 \rightarrow a, w_3 \rightarrow \varnothing, w_4 \rightarrow \varnothing$; <u>book b</u>: $w_1 \rightarrow \varnothing, w_2 \rightarrow b, w_3 \rightarrow \varnothing, w_4 \rightarrow b$

It is sometimes thought that English sentences like (8a) do not have a reading in which the permission extends to the full set of books. I believe this is incorrect. If one utters (8a) and Bill reads all the books, one can hardly say that he has been disobedient. The present account allows for this, as witnessed by w_2 in the model illustrated in (8c).

We see that in the case of necessity, fluctuation is violated: $\{<w_1, \{a,b\}>, <w_2, \{a,b\}>, <w_3, \{a,b\}>\}$. Consequently, *any* is unacceptable:

(9) a. *Bill must read any of these books.

b. $\forall x \ [x \leq_{w\text{-}a} \ \iota y \ [book(w_a)(y)] \rightarrow \forall w'.w_a \leq w'[read'(w')(x)(b)]]$

c. <u>book a</u>: $w_1 \rightarrow a, w_2 \rightarrow a, w_3 \rightarrow a$; <u>book b</u>: $w_1 \rightarrow b, w_2 \rightarrow b, w_3 \rightarrow b$

We have seen that a universal modal interacting with a universal quantifier grounded by the inner DP rules out the possibility of fluctuation while an existential modal allows for it. Thus fluctuation accounts for the distribution of

---

[8] The modal base in the cases under discussion would be restricted to worlds where the set of books includes at least a and b. Other books, if present, will not affect the computation of fluctuation (see ft. 5).

any in a context where no widening is possible.[9]   We now turn to episodic contexts.

6. Deriving FC *Any*: **Episodic Statements**

Unmodified *any* and partitive *any* are both unacceptable in episodic contexts. This follows straightforwardly if we make the standard assumption that universals presuppose non-empty domains of quantification. Whether the world variable on the DP is free or bound, fluctuation will be violated. An episodic statement refers to exactly one world. A single set of books, namely all the books in that world, was read in that world:

(10)a *Bill read any book / any of these books.

　b. $\forall x$ [book'$(w_a)(x) \rightarrow$ read'$(w_a)(x)(b)$]]

　c. $\forall x$ [$x \leq \iota y$ [book'$(w_a)(y)$] $\rightarrow$ read'$(w_a)(x)(b)$]]

There is no reference to possible individuals in this explanation as there was in Dayal (1998). The problem, rather, is the opposite. *Any* is ruled out because its denotation is too grounded in actuality, in the set of individuals in the actual world.

　　We turn now to the challenge of pining down the role of subtrigging. Note, first of all, that a garden variety account of modification does not help. As shown below, fluctuation remains elusive under modification (11a). Variation in the set of books needs to be derived via the relative clause in order to get the desired result (11b):

(11)a. $\forall x$ [[book'$(w_a)(x)$ & found'$(w_a)(x)(b)$] $\rightarrow$ read'$(w_a)(x)(b)$]

　b. $\forall x$ [[book'$(w_a)(x)$ & $\exists w'$. $w_a \leq w'$ [found'$(w')(x)(b)$]] $\rightarrow$ read'$(w_a)(x)(b)$]
　$\equiv$

　b'. $\forall x \forall w'$. $w_a \leq w'$ [[book'$(w_a)(x)$& found'$(w')(x)(b)$] $\rightarrow$ read'$(w_a)(x)(b)$]

　　The ameliorating effects of subtrigging is only observed with post-nominal phrasal modifiers (see Dayal 1995a and 1998). We can take this to show the crucial role of the world/situation variable, which prenominal modifiers lack. Also, subtrigging is felicitous in precisely those contexts where the speaker does

---

[9] Fluctuation may also explain why *n* in a sentence like *you may read any n of these books* can be any number less than the total number of books. If *n* = the number of books, *all n* is needed.

not have immediate knowledge of the full set of individuals involved. We may attribute the licensing of *any* by a modifier to its ability to introduce fluctuation cued to epistemic modality.

Indirect evidence for this approach comes from noting a subtle but clear shift in the domains of quantification of an unmodified universal quantifier and a modified universal quantifier. (12a) is an outright contradiction but (12b)-(12c) are not:

(12) a # I read every book but there were no books so I read nothing.

    b. I read every book I found but since I found no books, I read nothing.

    c. I read every book on the list but since there were no books listed I read none.

While an unmodified universal presupposes a non-empty domain of relevant entities, a modified universal lets in the possibility of variation down to the empty set. The claim here is that subtrigging rescues *any* because of this aspect of the semantics of modification. The speaker does not know how many books are at issue, the relevant set may have ten or five or four or zero members.

Further indirect support for this comes from considering mood and aspectual distinctions. In Dayal (1998) I noted that iterative contexts are conducive environments for subtrigged *any*, but iterativity is not essential (13a)-(13b). (13a), for example, is acceptable if uttered in a context where the relevant set of soldiers is not contextually salient. It has been noted by Quer (1998), on the basis of Catalan examples like (13a), that subtrigging is possible with subjunctive, not indicative, mood. This also holds for Spanish (Carlo Linares p.c.) and Italian (Chierchia 2006).[10] In an account of subtrigging in which the modifier provides a spatio-temporal anchoring to the world of evalutation, this is somewhat surprising – perhaps the opposite would be expected. From that point of view, the present proposal resonates better with cross-linguistic evidence:

(13) a. At the end of his speech, the President thanked any soldiers who had fought in the Gulf War.

    b. Anyone who was at the rally signed the petition.

(14) If every semanticist owned a villa in Tuscany, what a joy it would be.

Another piece of suggestive evidence comes from data like (14) from Percus (1997). *Every* in conditionals can have a non-fluctuating reading in which happiness depends on the semanticists in the actual world owning villas in

---

[10] Menendez-Benito (2005) considers subtrigging marginal or unacceptable in Spanish. Clearly, there is some variation in judgments. I base the discussion here on the judgments I have elicited.

counterfactual worlds. It can also have a fluctuating reading where the semanticists in the counterfactual worlds are villa owners. If the present account is on the right track, a FCI is predicted to have only the fluctuating reading. This prediction is borne out. (14) with *any* has only the fluctuating reading and the same holds for the Spanish counterpart with FCI.[11]

Let me reiterate that the relevant aspect of modification that makes FC *any* acceptable in episodic statements is a type of modality that allows fluctuation.[12] That is, the possible worlds that are invoked vary in the set of relevant entities, not just in the identity of a fixed set of entities. The importance of this will become clear in our discussion of variation among FCI. I should note here that the modality of subtrigging has previously been suggested by other researchers. The present proposal, though different in details of implementation, is in line with them rather than with Dayal (1998) where the vagueness requirement was used to capture the perceived indeterminacy.

7. Deriving FC *Any*: Negative Episodic Statements

Since negation does not introduce reference to other worlds, it is predicted that FC *any* would be ruled out in negative episodic statements as well, a prediction that seems to be counter-exemplified by sentences like (15a). It readily allows for a reading in which the universal takes scope over negation. It is also possible to get the other scope order, with the addition of *just* and/or with special intonation, as in (15b). The problem can be articulated more clearly in a language like Italian where FCI and NPI are lexically distinct. (16a)-(16b) are from Chierchia (2006):

(15) a. Bill didn't read any book.
  b. Bill didn't read (just) any$_F$ book, he read *Remembrance of Things Past*.

(16) a. Non leggerò qualunque libro
    (I) won't read (just) any book.
  b. Non leggerò qualunque libro che mi consiglierà Gianni
    (I) won't read any book that Gianni will recommend to me.

Unmodified FCI *qualunque N* when accompanied with special intonation has only a $\neg\forall$ reading. Without special intonation, subtrigging is required in order

---

[11] The English sentence would also have an existential reading, roughly *if even one semanticist had a villa in Tuscany what a joy it would be*. This reading I take to be due to NPI *any*.

[12] Cross-linguistically, this modality may be covert or encoded in mood/aspect morphology.

Variation in English Free Choice Items

for it to be acceptable. The reading it has in this case is $\forall\neg$. This suggests that the $\forall\neg$ reading of the non subtrigged English (15a) may, in fact, be the $\neg\exists$ reading of NPI *any*.

Focusing on Italian, then, a non-subtrigged statement like (16a) is not acceptable since it is an episodic statement that does not admit fluctuation. Once it is subtrigged, a negative statement like (16b) becomes acceptable, just like its affirmative counterpart. In the non-subtrigged case, the set of books that will be unread will be the full set in the actual world. In the case of subtrigging, the set of books unread will vary depending on what Gianni might recommend.

Turning now to the $\neg\forall$ reading, the primary question has to do with the role of intonation in promoting it. In accounting for this reading, I will look at examples like (17) in which *just any* goes hand in hand with supplemental *any*, a term due to Jennings (1994). These are cases in which *any* is linked to an indefinite. This can be overt (17a) or covert (17b): *press a key, any key*. Note that *just any*, like supplemental *any*, is not improved by subtrigging – special intonation is still needed to make it acceptable. Typically, the use of *just any* in affirmative statements like (18), an example fashioned after one by Vlachou (2007), also seems to require an additional clause indicating a purpose of some kind (see Dayal 2004 for some discussion of this):

(17) a. He read a book, not just any book.
    b. (To continue), press any key / any of these keys.

(18) Bill picked up a book, (just) any book, and walked out of the store.
    $\exists x [book(w_a)(x) \& pick\text{-}up (w_a)(x)(b)]$ *Indefinite antecedent*
    $\forall y [book(w_a)(x) \rightarrow \exists w' w_a \leq w' [pick\text{-}up(w')(x)(b)]]$ *Supplemental Any*

Restricting the modal base to worlds where Bill's purpose in picking up the book is fulfilled, we can take the semantic contribution of supplemental *any* to give rise to the (conversational) implicature that no individual who has the nominal property is special with respect to the verbal property. That is, there is nothing special about the book Bill picked up. My claim is that when there is negation in the sentence, it targets this aspect of the meaning, leaving the indefinite antecedent untouched. Bill read a book that was special, namely *Remembrance Of Things Past*. This is, then, a case of what is generally classified as metalinguistic negation, which is known to be signaled by intonation:

I should acknowledge here that while the $\forall\neg$ reading of FCI are well accounted for under the present approach, the $\neg\forall$ readings require further study. To conclude our current discussion of FC *any*, I have presented a uniform account of two recalcitrant problems – subtrigging and the partitive puzzle – and shown its usefulness in accounting for the occurrence of FC *any* in negative episodic statements as well.[13] I will now turn to the question of variation among FCI in English.

8. Variation in English FCI: Quantification

English has three FC items, *any*, *wh+ever* and *some N or other*. Of these, the first two have received considerable scrutiny but insights from one, unfortunately, have not typically been used to understand the other. I would like to propose that all FCI involve some form of indeterminacy but they differ along two dimensions, quantificational force and the strength of the indeterminacy they demand. Let us look at the quantificational dimension first.

In the case of *any*, I have argued here, as in earlier work, that it is a universal. Free relatives have been independently shown to be definites (Jacobson 1995, Dayal 1995b, 1997).[14] And *some N or other* appears, on the face of it, to be an indefinite. Nevertheless, I will repeat some of the diagnostics that establish these correlations.

(19a), like its plain counterpart or an ordinary definite, entails maximality (19b) and can be used anaphorically (19c).[15] By contrast, *some N or other*, like its plain counterpart, does not imply maximality (20b) and cannot refer anaphorically (20c):

(19) a. John read whichever book(s) Bill bought.
   b. #John read whichever books/the books Bill bought but not every book.
   c. Bill bought some book(s). John read whichever book(s)/the book(s) Bill bought.

(20) a. John read some book or other that Bill bought.
   b. John read some book or other that Bill bought but not every book he bought.

---

[13] The application of this approach to FCI in comparative clauses discussed by Zepter (2003) and in the canasta examples discussed by Menendez-Benito (2005), I leave for future research.

[14] For interesting cases of existential free relatives, see Caponigro (2003).

[15] I do not give the plain free relative counterpart with *which book(s)* here since English does not allow internal heads without *–ever* but the version with *what* could also make the case.

 c. #Bill bought some book. John read some book or other that Bill bought.

Finally, *-ever* free relatives can be distinguished from FC *any* by examining partitive readings. FC *any*, like the regular universal, has a strictly distributive reading. (21a) can only be true if John reads two thirds of each book. An *-ever* free relative or a definite has an additional collective reading: John's reading six out of nine books completely and not touching the other three is enough to make (21b) true:

(21) a. John has read two thirds of any book(s)/every book Bill bought.
  b. John has read two thirds of whichever books/the books Bill bought.

As we will see, some attention has been paid in the literature to differences between "universal" and "existential" FCI but not enough has been done to separate "definite" from "universal" FCI.   The data here underscore the importance of doing so.[16] Under the current approach, where the FCI attaches to ordinary quantificational expressions, these differences are transparently reflected.

9. Variation in English FCI: Strong vs. Weak Indeterminacy

Let us start our discussion of indeterminacy with some data involving FC *any*. *Any* cannot be used to answer questions requiring full specification. (23), unlike (22), seems to reject the premise of the question, as do (24b')-(25c):

(22) Speaker A: Which books did Bill read?
   Speaker B: He read every book he found.
   Speaker A: Yes, but I want to know exactly which books.
   Speaker B: Oh, I see. That would be: Namesake, Catcher in the Rye…

(23) Speaker A: Which books did Bill read?
   Speaker B: He read any book he found.
   Speaker A: Yes, but I want to know exactly which books.
   Speaker B: Oh, I couldn't tell you exactly which ones.

(24) Speaker A: Who wants sushi for dinner?
   Speaker B: Everyone I checked with wants sushi, Sue, Tim John, everyone.
  #Speaker B': Anyone I checked with wants sushi, Sue, Tim, John, anyone.

---

[16] Giannakidou and Cheng (2006) recognize definite FCI but do not distinguish it from universal FCI since they take items like FC *any* to be generic indefinites. Similarly, for Vlachou (2007).

Veneeta Dayal

(25) a. Who would like dessert?
    b. Everyone would like dessert.
    c. #Anyone would like dessert.

   -*ever* free relatives often appear to be synonymous to statements with FC *any*. In addition, they have an additional "identity" reading, prominent in episodic statements (26). In Dayal (1997) I argued for a requirement of indeterminacy with respect to identity as basic to *wh-ever* free relatives. I include *some N or other* for completeness:

(26) a.   Bill eats whatever / anything Sue cooks.
    b. Bill was eating whatever Sue had cooked.

(27) a. What she is cooking, namely ratatouille, uses onions.
    b. *Whatever she is cooking, namely ratatouille, uses onions.

(28) a.   I read some book or other, I can't remember it's name.
    b. *I read some book or other, namely *None to Accompany Me*.

To capture this property of –*ever* I built into its meaning universal quantification over worlds that are i(dentity)-alternatives to each other.[17] [18] I claimed that this was sufficient to account for both readings. The free choice reading of –*ever* free relative would be derived when the identity reading combined with generic tense-aspect.[19]

   To sum up, a distinction between two types of indeterminacy is needed to handle the range of English FCI. The indeterminacy of FC *any* is strong in the sense that the identity of the relevant set is not known to the speaker, because it is, in principle, not knowable: the set varies across worlds. The indeterminacy of -*ever* and *some N or other*, is weak in that it only requires ignorance about the identity of the set.

---

[17] Typically the identity reading is illustrated in terms of (in)ability to name an object but this is a simplification. Dayal (1997) mentions –*ever* free relatives like *John Smith/That person, whoever he is, has a problem* where knowledge of someone's name or ability to point to someone still allows for indeterminacy. See Heller and Wolter (2008) for an insightful analysis.
[18] Indeterminacy does not hold in this clear-cut a form when the free relative is plural.
[19] von Fintel (2000) attributes to Anna Szabolcsi the observation that there are languages like Hungarian in which separate items are used for identity and FC readings.   This would prompt a modification of this approach, either for those languages or for free relatives more generally.

10. The Indeterminacy of *wh+ever* and *some N or other*

Following our general approach so far, the denotation of an *-ever* free relative will be the same as that of a plain free relative. It will denote the maximal entity with the relevant property. Its conventional implicature will convey that there are several i-alternatives for that referent. This is, in fact, what von Fintel (2000) proposes:

(29) wh+ever⟧ = λP ιx [P (w) (x)]   *maximality*

    ∃w' ∃w'' ∈ F: ιx [P (w') (x)] ≠ ιx [P (w'') (x)] *ignorance implicature*

von Fintel has the *-ever* free relative trigger the presupposition that among the worlds in the modal base F, supplied by context, there is variation as to the identity of the referent.[20] That is, the notion of i-alternatives is cast in terms of epistemic modality and couched as a presupposition. In the case of (26b) for example, the assertion is simply that of a regular free relative, namely that Bill was eating the thing that Sue had cooked. But it is presupposed that the speaker doesn't know the identity of what she had cooked.

    von Fintel also discusses what he calls the indifference reading of *–ever* free relatives (see also Tredinnick 2005). In examples like (30a) the speaker may well be aware of the identity of the object but may not care about the identity per se. Its being handy is what counts. von Fintel proposes the presupposition in (30b) to capture this:

(30) a. I grabbed whatever tool was handy.

    b. ∀w' ∈ min_w [F∩ (λw' [ιx [P(w')(x)] ≠ ιx [P(w)(x)]]) :
                                Q(w')(ιx [P(w')(x)]) = Q(w) (ιx [P(w)(x)])
                                        indifference implicature

Regardless of what the free relative denotes, the truth of the statement remains constant across the worlds of the modal base that are minimally different from the actual world. The *min*-operator is taken to be the trigger for an existential presupposition, ensuring that there are varying worlds. I refer the reader to von Fintel's discussion for further details and for possible ways of unifying the ignorance and indifference readings. The point of relevance for us is that the assertive component captures the quantificational force of *–ever* free relatives and the conventional implicature its weak indeterminacy.

---

[20] Ignorance can be cued to the hearer, though it is usually cued to the speaker (von Fintel 2000).

Another significant advance in our understanding of *–ever* free relatives is due to Heller and Wolter (2008). Focusing on the ignorance reading of *–ever* free relatives, they present a nuanced discussion of the relationship between rigidity and identity. They argue for a fundamental shift in the modeling of these concepts, adopting Gupta (1980). The significance of their contribution is in drawing attention to the critical role of the internal head of the relative clause in determining possibilities for reference and in using the notion of sorts and trans-world identity to explain the relation between these possibilities and indeterminacy. Again, I refer the reader to the original article for details. I believe that it should not be difficult to incorporate their insights into a multi-dimensional approach to the meaning of *–ever* free relatives, allowing for maximality in the truth-conditional component and weak indeterminacy in its implicational component.

To complete the picture, I propose that *some N or other* differs from *-ever* in having existential quantificational force while sharing the same conventional implicature. Both expressions, because they only require weak indeterminacy, are therefore compatible with episodic statements. Instead, FC *any* shows the sensitivity to modality that it does because of its requirement of strong indeterminacy.

11. FC *any:* Compositionality and Scope

I have accounted for variation along two dimensions, quantification and indeterminacy. As indicated, dividing up the meaning of *-ever* free relatives into a quantificational component and an implicature is in line with current thinking on the topic. In the case of FC *any*, however, the present ideas do not mesh with a significant body of work that I do not have space to discuss properly, namely Kratzer and Shimoyama (2002), Chierchia (2006) and Fox (2007), among others. These articles luckily are well known enough that I can assume familiarity with the essentials of the approach they advocate and settle for listing some of their key contributions. Kratzer and Shimoyama, for example, provide a notion of strengthening that reconciles an indefinite FCI over a widened domain with affirmative contexts. Chierchia extends this idea to allow for universal FCI as well as NPI, giving a principled explanation for the fact that many languages have the same lexical item for both while as many do not. He also frames his theory of polarity sensitive items within a framework of recursive pragmatics, as does Fox in his account of FC effects with disjunction. This, then, is an approach with an impressively wide reach.

Variation in English Free Choice Items

Compelling as it is, there remain open some crucial questions from the perspective of the issues discussed here: differences between definite *–ever* free relatives and FC *any*, and the distribution of partitive *any*. It is not clear to me, for example, where in Chierchia's system a three-way quantificational distinction could be drawn and I do not see a solution for the partitive problem. The approach, in my view, suffers from an over-reliance on the nominal property inherent to the notion of widening. The distributional facts we have looked at call for a more pliable conceptual tool. The proposed implicature of fluctuation has this pliability, admitting widening for generic *any,* narrowing for subtrigged *any*, and crucially, modal manipulation for partitive *any*.

That said, the fluctuation-based approach remains incomplete. It is silent, for example, on the precise character of the conventional implicature, the relationship between FCI and NPI, on the relationship between FCI and disjunction, and the historical connection between indefinites and FCI/NPI. In the remaining space I will make some comments on one of these questions, namely the type of multi-dimensional meaning I have ascribed to FC *any,* indicating the direction in which I hope to take this work in the future.[21]

I propose the slightly revised meaning in (35a) for FC *any*. Its primary meaning differs from the regular $\forall$ in having an extra variable Z, of type $<<s,t>,t>$, with scope over the VP meaning. (35b) is as before, except for the abstraction over properties P and Q. The two dimensions of meaning can now be computed compositionally. Consider the LF in (36b), with the modal adjoined to TP by QR.[22] (37a) shows how the primary meaning of *any*, given in (35a) figures in the computation. The main novelty, due to the revision in (35a) is that there remains a variable inside the DP even after it combines with the VP. This reverses the usual order of composition of a modal with the clause:

(35) a. Any⟧ $=\lambda P\,\lambda Q\,\lambda Z\,\forall x\,[P(w)(x) \rightarrow Z\,\lambda w'\,Q(w')(x)]$ *Modal $\forall$ Quantifier*

b. $\lambda P\,\lambda Q\,\neg\,\exists X\,\forall w'\,w_a \leq w'\,\lambda x\,[P(w')(x)\,\&\,Q(w')(x)] = X$ *F-implicature*

---

[21] To account for the prevalence of languages with the same lexical item for NPI/FCI, perhaps a homogeneity condition uniting all FCI and also FCI and NPI is needed (Jackson 1995, Dayal 1998 and Zepter 2003): $\neg\,x\quad y\,[P(y)\,\&\,P(x)\,\&\,x{\neq}y\,\&\,Q(x)\,\&\,\neg Q(y)]$. This would say, in effect, that statements in which such items occur must have universal import. This may also explain why neutral $\neg\forall$ readings are not available for FC *any* or for *-ever* free relatives. But to have bite, the application of this condition has to be appropriately characterized and modulated.

[22] Some adjustments will likely be needed to account for generic and episodic cases.

Veneeta Dayal

(36) a. Any of these students can / *must win.

   b. $[_{TP'}$ can/*must $[_{TP} [_{DP}$ any of these students] $[_{VP}$ win]]]

(37) a. $⟦DP⟧ = \lambda Q \lambda Z \forall x [x \leq_{w-a} \iota y [students(w_a)(y)] \rightarrow Z \lambda w' Q(w')(x)]$

   $⟦VP⟧ = \lambda w' \lambda x [win(w')(x)]$

   $⟦IP⟧ = ⟦DP⟧ (⟦VP⟧) =$

   $\lambda Z \forall x [x \leq_{w-a} \iota y [students (w_a) (y)] \rightarrow Z \lambda w' win (w')(x)]$

   $⟦◇⟧ = \lambda p \exists w' w_a \leq w' [p(w')] ; \quad ⟦□⟧ = \lambda p \forall w' w_a \leq w' [p(w')]$

   $⟦TP'⟧ = ⟦TP⟧ (⟦◇⟧ / ⟦□⟧) =$

   $\forall x [x \leq_{w-a} \iota y [students (w_a) (y)] \rightarrow \exists w' w_a \leq w' [win (w')(x)]]$
   for 'can'

   $\forall x [x \leq_{w-a} \iota y [students (w_a) (y)] \rightarrow \forall w' w_a \leq w' [win (w')(x)]]$ *for 'must'*

(37) b. $⟦DP⟧ = \lambda Q \neg \exists X \forall w' w_a \leq w' \lambda x [x \leq_{w'} \iota y [students(w_a)(y)] \&$

   $Q(w')(x)] = X$

   $⟦VP⟧ = \lambda w' \lambda x [win(w')(x)]$

   $⟦IP⟧ = ⟦DP⟧ (⟦VP⟧) =$
   $\neg \exists X \forall w' w_a \leq w' \lambda x [x \leq_{w'} \iota y [students (w_a)(y)] \&$
   $win(w')(x)] = X$

(38) $[_{TP''} [_{DP-i}$ any of these students] $[_{TP'}$ can $[_{TP} t_i [_{VP}$ win]]]]

Now let us compute the secondary meaning of *any*, given in (35a) with reference to (37b). Note that what we derive in (37b) is compatible with the primary meaning for possibility *can* in (37a) but not the primary meaning for necessity *must* in (36a). That is, it correctly predicts the grammaticality judgements for the sentences under discussion.

   Let us now consider the other LF for the sentence *Any of these students can win*, in (38). The problem is that the modal is in the wrong place for functional application to take place. Crucially, the secondary meaning (and the way we now have defined the primary meaning) has to be computed without the modal being part of the property Q. So, it seems that the only interpretable derivation for the sentence is one in which $\forall$ c-commands $◇$. However, given the way we

have now defined the primary meaning of *any,* we derive the interpretation in which $\diamond$ ends up having scope over $\forall$, in line with conventional wisdom about the scopal properties of FC *any*.[23][24]

Finally, let us consider the formal status of the requirement of fluctuation. That it is not part of the assertion is shown by the fact that *any of these students must win* is not false, as would be expected if fluctuation were part of the assertion. But whether fluctuation is a presupposition or a conventional implicature of a different kind (see Potts 2005) remains to be established.

I hope to have shown in this paper that the line of thinking I am pursuing has some initial merit. As I have indicated there remain many issues to be settled. Unfortunately, they must remain promissory notes for the present.

References

Caponigro, Ivano. 2003. *Free Not to Ask: On the Semantics of Free Relatives and Wh-Words Cross-linguistically*, Doctoral Dissertation, UCLA.

Carlson, Greg. 1981. Distribution of free choice *any*. *Proceedings of CLS* 17.

Chierchia, Gennaro. 2006. Broaden your views: Implicatures of domain widening. *Linguistic Inquiry* 37:535-590.

Dayal, Veneeta. 1995a. Licensing *any* in non-negative/non-modal contexts. *Proceedings of SALT V*.

Dayal, Veneeta. 1995b. Quantification in correlatives. In *Quantification in Natural Language* eds. Emmon Bach, Eloise Jelinek, Angelika Kratzer and Barbara Partee, 179-205. Dordrecht: Kluwer.

Dayal, Veneeta. 1997. Free relatives and *–ever*: Identity and free choice readings. *Proceedings of SALT VII*.

Dayal, Veneeta. 1998. *Any* as inherently modal. *Linguistics and Philosophy* 21:433-76.

Dayal, Veneeta. 2004. The universal force of free choice *any. Linguistic Variation Yearbook* 4:5-40.

von Fintel, Kai. 2000. Whatever. *Proceedings of SALT X*

Fox, Danny. 2007. Free choice disjunction and the theory of scalar implicatures. In *Presupposition and Implicature in Compositional Semantics* eds. Uli Sauerland and Penka Stateva. 71-120. New York: Palgrave Macmillan**.**

Giannakidou, Anastasia, and Lisa Cheng. 2006. (In)definiteness, polarity, and the role of wh- morphology in free choice. *Journal of Semantics* 23:135-183.

---

[23] Though different in details, my proposal here is reminiscent of the analysis in Saeboe (2001), where FCI induces an intensional interpretation.

[24] Note that FC *any* can take scope under other quantifiers: *Every student read any books on giraffes he found* (Carlson 1981). This does not seem problematic for the present account.

Gupta, Anil. 1980. *The logic of common nouns: An investigation in quantified modal logic*. New Haven: Yale University Press.

Heller, Daphna, and Lynsey Wolter. 2008. Identity and indeterminacy in –*ever* free relatives. *Proceedings of SALT 18*.

Horn, Lawrence R. 2001. Any and –ever: free choice and free relatives. *Proceedings of the 15* [th] *Annual Conference of the Israeli Association for Theoretical Linguistics* 71-111.

Jackson, Eric. 1995. Negative polarity and general statements. *Proceedings of SALT V*.

Jacobson, Pauline. 1995. On the quantificational force of English free relatives. In *Quantification in Natural Language* eds. Emmon Bach, Eloise Jelinek, Angelika Kratzer and Barbara Partee. 451-86. Dordrecht: Kluwer.

Jennings, Raymond E. 1994. *The genealogy of disjunction*. New York: Oxford University Press.

Kadmon, Nirit, and Fred Landman. 1993. Any. *Linguistics and Philosophy* 16. 353-422.

Kratzer, Angelika, and Junko Shimoyama. 2002. Indeterminate pronouns: the view from Japanese. *Paper presented at the 3* [rd] *Tokyo Conference on Psycholinguistics*.

LeGrand, Jean E. 1975. *Or and Any: The Semantics and Syntax of Two Logical Operators*. Doctoral dissertation, University of Chicago.

Menendez-Benito, Paula. 2005. *The Grammar of Choice*, Doctoral dissertation, University of Massachusetts, Amherst.

Percus, Orin. 1997. *Aspects of 'A'*. Doctoral dissertation, Massachusetts Institute of Technology.

Potts, Christopher. 2005. *The Logic of Conventional Implicatures*. Oxford: Oxford University Press.

Quer, Josep. 1998. *Mood at the Interface*. Doctoral dissertation, University of Utrecht. The Hague: Holland Academic Graphics.

Saeboe, Kjell Johan. 2001. The semantics of Scandinavian free choice items. *Linguistics and Philosophy* 24. 737-787.

Treddinick, Victoria. 2005. On the semantics of free relatives with -ever, Doctoral dissertation, University of Pennsylvania.

Vlachou, Evangelia. 2007. *Free choice in and out of context: semantics and distribution of French, Greek and English Free Choice Items*. Doctoral dissertation. University of Sorbonne-Paris IV and University of Utrecht.

Zepter, Alex. 2003. How to be Universal when you are Existential: Entailment along a Scale. *Journal of Semantics* 20:193-237.

Discourse Particles in Questions

Josef Bayer
Universität Konstanz, Germany

Abstract

Discourse particle play an important role in the formation of utterances. They signal the relation of speaker and hearer with respect to the propositional content at issue. For linguists who want to attain a deeper understanding of the syntax to discourse mapping and clausal architecture in general, the study of discourse particles can provide a new window.The present article covers some ground in the limited domain of wh-questions in German.

1. Introduction[1]

Discourse particles are elements which play a role when sentence types (declaratives, interrogatives, imperatives, optatives etc.) connect to discourse. They make a noticeable semantic contribution which, however, concerns "expressive" rather than "propositional" meaning (cf. Potts, 2005), as already noticed by Georg von der Gabelentz (1891): *When I insert* [these particles, JB] *into my speech, the reason for doing so cannot be immediately found in the subject matter of my speech but rather in an emotional need of the speaker*.

German abounds with discourse particles (a.k.a. "modal" particles) but they can be found in many other languages or language families as well, e.g. in Chinese (Paul, 2008) and in Indo-Aryan languages (Dasgupta, 1984; 1987).

Here we concentrate on some discourse particles in German questions and what can be derived from their properties about clausal architecture and syntactic processes. The article is organized as follows: Section 2 provides some general background on the nature of discourse particles. Section 3 contains an analysis of discourse particles in German wh-questions which will start with semantics and then go over to various syntactic issues of word order and phrase structure. The core proposal is a structure in which the particle is a functional head which selects VP. Section 4 turns to a more marked construction in which the particle forms a constituent with the wh-phrase. This leads to reconsidering the categorial

---

[1] A fuller version of this article has been prepared in Bayer & Obenauer (submitted).

Josef Bayer

nature of particles in section 5. Section 6 shows how the two constructions interact. Section 7 contains a conclusion.

2. Some General Background

Without trying to go into any details I would like to present here a list of properties that have been found to apply to discourse particles in German. Some of these properties are rigid across all vocabulary items, others are subject to some variation. According to what can be found in the literature, discourse particles ...

    i.   mostly have a counterpart to which they are historically related
   ii.   are the result of grammaticalization
  iii.   are often adverbial in nature but are distinct from adverbs
   iv.   are immobile
    v.   can usually not bear stress
   vi.   cannot be coordinated
  vii.   cannot be used in isolation, e.g. as answers
 viii.   are in their typical occurrences mono-syllabic
   ix.   are modificational, i.e. are mostly optional
    x.   can to a certain extent appear simultaneously in the clause
   xi.   appear in fixed order (comparable to the order of adverbs)
  xii.   are sensitive to sentences types (assertive, ±wh interrogative, imperative etc.)
 xiii.   belong to „expressive" rather than „propositional" meaning
 xiv.   appear – due to xiii – generally in root-clauses
  xv.   can arise in non-root clauses only under special conditions

Here we will concentrate on four particles which appear in wh-questions: *denn* ("then"), *nur* ("only") – and its near-equivalent *bloß* ("barely") – and *schon* ("already"). The translations in brackets are literal and should not suggest that this meaning enters semantic composition in the formation of utterances with discourse particles necessarily. Section 3.1. will turn to this question.

3. Discourse Particles in German Questions

Independent of the particular language, the pertinent questions are the following: (i) What is the function of discourse particles in clause structure? (ii) How can their role in the shaping of the illocutionary force of an utterance be accounted for?

3.1 Semantic Contribution

In each case, the particle adds a certain so-called "expressive" meaning to the question. Researchers of all kinds of persuasions agree that these particles affect the level of utterance (illocutionary force) rather than the propositional level (including "at-issue entailments", Potts, 2005: 156). Semantic features have been proposed by Thurmair (1989), most of which induce a special context dependence (to previous discourse or to the state of knowledge of the interlocutors). In (1)-(3) the semantic contribution of the particles *denn* (obviously connected to Engl. *then*), *nur/bloß* ("only") and *schon* ("already") is indicated in brackets:

(1) Wo   habe ich **denn**   meine Schlüssel hingelegt?
     where have I   DENN my   keys   put-down
     "Where did I put my keys (I'm wondering)?"

(2) Wo   habe ich **nur** / **bloß**   meine Schlüssel hingelegt?
     where have I   NUR / BLOSS my   keys   put-down
     "Where did I put my keys (I have already looked everywhere)?"

(3) Wo   ist er **schon**   gewesen?
     where is he SCHON   been
     "Where has he been?" (meaning: He hasn't been in important or desirable places)

In each case, deletion of the particle would turn the utterance into a straight information seeking question without the special semantic twists indicated in braces.

- *Denn* is most general. In an information seeking question it demands that the expected true answer p updates the common knowledge of speaker and addressee in such a way that p is relevant to the knowledge of the speaker. *Denn* is notoriously infelicitous in out-of-the-blue questions. If I meet a stranger for the first time I can hardly ask him/her *Wie spät ist es denn?* ("How late is it then?") because *denn* does not make reference to any piece of common ground of speaker and hearer.

- *Nur* (also *bloß*) is more special. It functions as an eliminative operator like in their function as a focus particle (e.g. *He drinks only TEA*, $\forall x$ [he drinks x $\rightarrow$ x = tea]). In wh-questions, it yields an interpretation according to which all constants which have so far been considered in the actual speech situation as being able to replace the wh-bound variable return the thruth value false. The

implicature is that the speaker has unsuccessfully tried to find the right value (cf. Obenauer, 2004 where wh-questions with this particle are called *I can't find the value questions*).

- *Schon* ("already") seems to be coined in analogy to its temporal use. Temporal "already" requires a scale on which *already* p denotes a state after $\neg$ p (*not-yet* p / *still* $\neg$ p). In wh-questions, *schon* may induce a similar scale by which the entities e that can replace the variable are ranked according to their plausibility of yielding a true answer, the implicature being that there are few or no e which are high enough on the scale to make the answer true. Wh-questions with *schon* are usually understood as rhetorical questions. In my view the interpretation in braces in (3) is not the semantic interpretation of the question but rather the implicated meaning.

3.2 Position Relative to Higher Adverbs

German root clauses showing the Verb-Second (V2) property, discourse particles occupy a high position below the landing site of the finite verb Fin° (related to C° in comp-clauses) and the higher (speech act, evaluative, evidential, epistemic etc.) adverbs. Reversal of the order yields a deviant result.

(4)   Wo   hat er **denn**   vermutlich seine Schlüssel hingelegt?
        where has he DENN presumably his   keys   put-down
        "Where did he presumably put his keys (I'm wondering)?"

        *vermutlich < **denn**

(5)   Wo   habe ich **nur** / **bloß**   dummerweise meine Schlüssel hingelegt?
        where have I   NUR / BLOSS stupidly   my   keys   put-down
        "Where did I stupidly put my keys (I have already looked everywhere)?"

        *dummerweise < **nur** / *dummerweise < **bloß**

Particles are obligatorily preceded by weak and clitic pronouns (cf. (6), (7)), and optionally preceded by topical elements (definite DPs, generic indefinites, stage setting adverbs and PPs etc.; cf. ((8) – (13)). These topical elements can also include contrastive topics.

(6)   Hat {mich/MICH} **denn** {*mich/MICH} jemand sprechen wollen?
        has me   DENN   me   someone speak   wanted
        "Did someone want to talk to me/ME?"

(7) Hat {es / 's } **denn** {*es / *'s } jemanden interessiert?
   has it   DENN it   someone interested
   "Did someone take an interest in it?"

(8) Wann könnte **nur** Otto den Brief ins Büro mitgenommen haben?
   when could   NUR Otto the letter to office along-taken   have
   "When could Otto have taken the letter to the office?"

(9) Wann könnte Otto **nur** ~~Otto~~ den Brief ins Büro mitgenommen haben?

(10) Wann könnte Otto den Brief **nur** ~~Otto~~ ~~den Brief~~ ins Büro mitgenommen haben?

(11) ?Wann könnte Otto den Brief ins Büro **nur** ~~Otto~~ ~~den Brief~~ ~~ins Büro~~ mitgenommen haben?

(12) Wo   kann ich **nur** einen Kaugummi   kaufen?
   where can   I   NUR a   chewing-gum buy
   Where can I buy chewing gum?

(13) *Wo kann ich einen Kaugummi **nur** ~~einen Kaugummi~~ kaufen?

In (12)/(13) we see that an indefinite cannot normally precede the particle. It must remain in VP where it is existentially interpreted. In those cases where it raises up, the indefinite receives a generic interpretation as in *Wo kann in dieser Welt ein Tiger nur ~~ein Tiger~~ überleben?* "Where on earth in this world could a tiger survive?

3.3 Stacking of Particles and Ordering

Discourse particles can be stacked, and if so they are stacked in fixed order (cf. Thurmair, 1987; Cogniglio, 2005). This is reminiscent of the ordering of adverbs studied in Cinque (1999). In wh-questions, *denn* – being the highest particle - precedes *nur/bloß* or *schon*.

(14) Wo   bist du **denn nur** / **bloß**   den ganzen Tag gewesen?
   where are   you DENN NUR / BLOSS the whole   day been
   "Where on earth have you been the entire day (I am wondering)?"

   *nur < denn   /   *nur < bloß

(15) Wo   wird er **denn   schon**   gewesen sein?
   where will he DENN SCHON been   be
   "Where will he have been?
   (meaning: He can't have been in important or desirable places)

   *schon < denn

Josef Bayer

 The following examples from the internet show combinations of *denn, wohl* (well) and *nur*.

(16) Er verstehe das ganze Schicksal überhaupt nicht, was er **denn wohl   nur**
he understand_{subj} the whole fate absolutely not what he DENNN WOHL NUR
verbrochen habe, **...**
commited   had_{subj}
"He would not understand this fate at all, (it being unclear) what kind of
   crime he should have committed"

*wohl < denn / *nur < wohl

(17) Wen kann er **denn wohl nur** mit "ganz schön dicht" gemeint haben?
who can   he DENN WOHL NUR with quite beautifully intoxicated meant   have
"Who can he have had in mind when he said "quite intoxicated"?

Notice that *denn+nur*, *denn+bloß* , *denn+schon*, *denn+wohl+nur* etc. cannot be analyzed as a "super particle". They also appear in non-adjacent positions, e.g. (14) could also be (18):

(18) Wo bist du **denn** den ganzen Tag **nur**/**bloß** ~~den ganzen Tag~~ gewesen?

3.4 Phrase Structure

A number of properties suggest that discourse particles are functional heads which precede VP/*v*P: Immobility, semantic bleaching, grammaticalization, phonological shape, and new evidence that will be given in §4 below. Discourse particles can be stacked, and discourse topics may move to a designated topic field to the left of the particle(s). The particles under closer consideration here, *denn*, *nur*, *bloß*, *schon*, arise in questions; they must be in the scope of an interrogative (Q- or wh-) feature in Fin°.

In fact, We assume here that in the German root clause, Fin° hosts the finite verb, and that in doing so it activates force. Thus, we assume here that V2-FinP = ForceP. Question-sensitive particles, abbreviated here as Prt_1 through Prt_n , appear as heads of phrases (PrtP) in the closest c-command domain of interrogative Fin° as shown in (19).
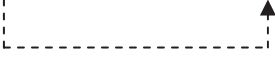
(19) $[_{FinP}$ (Wh) Fin° $[_{TopP}$ ... $[_{PrtP1}$ Prt_1° $[_{PrtP2}$ Prt_2° ... $[_{PrtPn}$ Prt_n° [ADV* $[_{VP/vP}$ ...]]]]] ...]]

3.5 Connecting to Force

The discourse particles (Prt) under consideration must occur in questions, and the questions must be root questions (i.e. questions with illocutionary force). This suggests roughly two options:

  (i) Prt is automatically part of the force system when merged to VP/vP.
       OR
  (ii) Prt accesses the force system via some operation: covert raising, move-F, or probe / goal agreement)

For reasons which follow shortly, we will argue in favor of (ii). Assume a sub-feature of force on Prt, <Prt Force>, by which agreement with interrogative force is established.

(20) [$_{FinP}$ (Wh) Fin° $_{<Interr>}$ [$_{TopP}$ … [$_{PrtP1}$ Prt° $_{<PrtForce>}$ [$_{VP/vP}$ …]]]]

With the force feature deleting in the base, agreement returns (20) as (21).

(21) [$_{FinP}$ (Wh) Fin° $_{<Interr, PrtForce >}$ [$_{TopP}$ … [$_{PrtP1}$ Prt° $_{<\text{PrtForce}>}$ [$_{VP/vP}$ …]]]]

Probe/goal agreement permits interrogative Fin to acquire the additional force value of a discourse particle. For reasons which cannot be laid out completely in this context, I assume that <PrtForce> establishes the link to the force system of the root calause while the lexical part of the particle stays behind and takes scope right where it has been merged.[2]

3.6 Discourse Particles in Dependent Clauses

In the majority of the data, discourse particles occur immediately in the root clause. Infinitival clauses do not seem to host particles at all, cf. (22) versus the ungrammatical (23). (24) does not count because it is monoclausal structure (signaled by { }) due to clause union.

(22) Wohin hast du **denn**   versucht [diesen Brief ~~wohin~~ zu schicken]?
      where have you DENN tried   this   letter   to send
      "Where did you try to send this letter (I am wondering)?"

(23) *Wohin hast du versucht [diesen Brief **denn** ~~wohin~~ zu schicken]?
       where have you tried   this   letter DENN   to send

---

[2] For details see Bayer and Obenauer (submitted).

Josef Bayer

(24) Wohin hast du diesen Brief **denn** ~~wohin~~ {zu schicken versucht}?
    where have you this   letter DENN   to send   tried

The reason could be that infinitives have a truncated functional structure which lacks PrtP.

    Finite complements do allow discourse particles under conditions which are revealed in the derivations below.

(25) a. Wie denkst du, dass es **denn**   weitergehen soll   mit euch?
        how think you that it DENN go-on   should with you
        "How do you think that the two of you should carry on?"

[datum from the internet]

    b. Wie denkst du [_{CP} ~~wie~~ dass es [_{PrtP} **denn** ~~wie~~ weitergehen soll mit euch]]?

(26) a. Wie denkst du dass seine Mutter **denn**   meint, dass es weitergehen soll
        how think you that his mother DENN thinks   that it go-on   should

    mit euch?
    with you

    b. Wie denkst du [_{CP} ~~wie~~ dass seine Mutter [_{PrtP} **denn** meint [_{CP} ~~wie~~ dass es ~~wie~~ weitergehen soll mit euch]]]?

If wh does not pass the particle, the particle is not licensed in the embedded clause:

(27) a. *Wem hast du erzählt dass Karl **denn**   recht hat?
        whom have you told   that Karl DENN right has
        "Who did you tell that Karl was right?"

    b. *Wem hast du ~~wem~~ erzählt [_{CP} dass Karl [_{PrtP} **denn** recht hat]]?

Approach (ii) of § 3.5 allows a natural explanation which is in harmony with cyclic wh-movement as well as with the idea of derivation by phase. Informally speaking, once a CP is built, Fin°_{<Interr>} can probe the locally available PrtP and acquire its <PrtForce> feature. Cyclic wh-movement will transport this feature outward to the locus of interpretable illocutionary force. As expected, occurences of different particles are also possible in different CP-cycles while the relative order, e.g. *denn* < *nur*, must be retained. Here is a possible and an impossible version of (25).

(28) a. Wie denkst du **denn**, dass es **nur** weitergehen soll   mit euch?
        how think you DENN that it NUR go-on   should with you
        "How do you think that the two of you should carry on?"

   b. *Wie denkst du **nur**, dass es **denn** weitergehen soll   mit euch?
     how think you NUR that it DENN go-on   should with you
     "How do you think that the two of you should carry on?"

PrtP can in principle be at an arbitrary distance from the root's force projection as long as it can make contact which it via derivation in phases.

4. Discourse Particles and Constituency

So far we have seen discourse particles in rigidly fixed pre-VP/*v*P position. Property iv. in § 2 says they are immobile. While this remains to be true, an important qualification must be added: The particles under consideration can also appear as a co-constituent of the wh-phrase and then move together with it, cf. (29) or the sluicing case in (30), – both data found on the internet.

(29) [Warum **bloß**]   ist ein Rauschenberg so teuer?
     why   BLOSS is a Rauschenberg so expensive
    "Why the hell is a (painting by) Rauschenberg so expensive?"

(30) Fran ist lustig   und erfolgreich ... und schwanger, aber [von wem **bloß**]!!
     Fran is humorous and successful   and pregnant but from who BLOSS
    "Fran is nice and successful .. and pregnant. But from who?"

Given the architecture in (19), how can we account for these cases? We exclude the possibility that wh moves and on its way "decapitates" the PrtP taking its head along. Notice there is no independently motivated process such as cliticization to the wh-phrase etc. Let us therefore assume that wh can separately be merged with Prt and project a "Small PrtP" which is then merged like any other constitutent, e.g. as a +wh DP, a wh PP, a wh AdvP etc. A wh-phrase that is merged with Prt bears extra heavy stress. After being merged with Prt, this wh-phrase moves to the left of Prt. The distinctive feature of the construction is the extra strong accent which signals emphasis.[3] We propose an implementation by which the relevant feature (emp) is an unvalued feature on the particle which attracts the emphatically accented wh for valuation.

(31) $Prt°_{uEmp}$ wh $_{Emp}$ $\Rightarrow$ [wh $_{Emp}$ [$Prt°_{uEmp}$ ~~wh~~ $_{Emp}$]]   "Small PrtP"

---

[3] Although emphasis requires extra accent, it needs to be distinguished from semantic focus or any other notion that relates to information structure. Behaghel (1932) speaks of accented topicalization in German as motivated by the speaker's affect or excitement. For further discussion cf. Bayer (2001).

Josef Bayer

The resulting Small PrtP is marked for emphasis; due to the fact that Prt has attracted a wh-phrase it is, of course, also +wh and undergoes the usual wh-movement.

Particles can be stacked in small PrtPs, in which case the particles retains exactly the same order as in pre-VP/vP position. The following examples are again taken from internet sources:

(32) [Warum **denn   nur**] kann AMD ihre CPUs billiger   anbieten als Intel?
    why   DENN NUR can   AMB its   CPUs cheaper offer   than Intel
    "Why on earth can AMD offer their CPUs cheaper than Intel

                            (I am wondering)"?

(33) [Wie **denn   bloß**]   kann ich sie fangen
    how DENN BLOSS can   I   her catch
    "How on earth can I catch her (I am wondering)"?

(34) [Wer **denn schon**] würde es der Meinung   eines Dritten verbieten wollen
    who DENN SCHON would it the opinion$_{dat}$ a   third$_{gen}$ prohibit   want

    die deine zu meiner zu machen?
    the your to mine   to make
    "Who would deny the opinion of a third person to make yóur opinion míne?"

These examples show that emp-valuation is recursively available in the sense that a single wh-phrase can value more than one emp-feature.

(35) a.   Prt1°$_{<uEmp>}$ wh $_{<Emp>}$ ⇒   Move wh
    b.   [wh $_{<Emp>}$ [Prt1°$_{<uEmp>}$ ~~wh~~ $_{<Emp>}$]]   ⇒   Merge Prt2°
    c.   Prt2°$_{<uEmp>}$ [wh $_{<Emp>}$ [Prt1°$_{<uEmp>}$ ~~wh~~ $_{<Emp>}$]]   ⇒ Move wh
    d.   [wh $_{<Emp>}$ [Prt2°$_{<uEmp>}$ [~~wh~~ $_{<Emp>}$ [Prt1°$_{<uEmp>}$ ~~wh~~ $_{<Emp>}$]]]]

Given that discourse particles affect the force of the sentence, which I assume presupposes propositional scope, the small PrtP seems to be notoriously unfit to meet this requirement. This problem can naturally be resolved if we assume that the small PrtP maps onto the "big" pre-VP/vP PrtP as suggested in (19) and following representations, and that this happens in analogy to wh-movement. According to this idea, the +wh small PrtP first moves to the left edge of VP/vP where it values a the feature <uPrt> of a silent Prt-head and activates the scope of Prt. In the next step the +wh small PrtP moves on to the left edge of FinP where it values the <uInterr> head associated with the Fin-head in V2 position. The derivation, in which the step of wh-movement is abbreviated for the sake of readability, runs as in (36).

(36) a. $[_{VP/vP} ... [wh Prt] ...]$ $\Rightarrow$ Merge a silent Prt
  b. $[_{PrtP} Prt°_{<uPrt>} [_{VP/vP} ... [wh Prt] ...]]$ $\Rightarrow$ Move [wh Prt]
  c. $[_{PrtP} [wh Prt] Prt°_{<\text{uPrt}>} [_{VP/vP} ... [\text{wh Prt}] ...]]$ $\Rightarrow$ Move [wh Prt]
  d. $[_{FinP} [wh Prt] Fin°_{<\text{uInterr}>} [_{PrtP} [\text{wh Prt}] Prt°_{<\text{uPrt}>} [_{VP/vP} ... [\text{wh Prt}] ...]]]$

 Speaking in terms of X-bar theory and Rizzi's criterion approach to functional heads, the small +wh PrtP runs through the specifier of a PrtP "identifying" Prt lexically, and then moves to SpecFinP (perhaps via other CP-cycles) "identifying" Fin° as a wh interrogative.

    Small PrtPs with stacked particles as seen in (32) through (34) can easily be integrated in this account because the small PrtP can pass through as many PrtPs as there may be built over VP/*v*P using the innermost Prt first. This reflects the parallel order of discourse particles that can be observed between the canonical architecture in (19) and the small PrtPs with stacked particles.

5. A Note on the Categorial Nature of Particles

The categorial status of discourse as well as focus particles is controversial. Bayer (1996; 1999) takes focus particles as heads which may select a non-propositional XP.[4] In a V2-language like German this appears to be close to inevitable because the combination particle plus an arbitrary XP appears in pre-verbal position as can be seen in the following examples.

(37) a. **Nur** HELENE hat mit   Klaus telefoniert
       only Helene   has with Klaus telephoned
      "Only Helene talked over the phone with Klaus"
   b. **Nur** MIT KLAUS hat Helene telefoniert
       only with Klaus   has Helene telephoned
       "Helene talked over the phone only with Klaus"

   c.   **Nur** TELEFONIERT hat Helene mit   Klaus
       only telephoned   has Helene with Klaus
       "Helene only telephoned with Klaus"
                      (she didn't to anything else with him)

---

[4] Rothstein (1991) suggests next to lexical and functional heads *minor functional heads* for functional elements that do not project categorial features. In this case, the category of the non-head is retained. Rothstein has certain degree words, particles like English *too* and conjunctions in mind. If focus and discourse particles are minor functional heads, they appear as syncategorematically introduced. This may be one of the reasons why many linguists have difficulty attributing head status to them and prefer to speak of adverbs.

Josef Bayer

Jacobs (1983) and more recently Büring & Hartmann (2001) nevertheless argue that focus particles are adverbs, and that as such they do not form a constituent with XP. According to their theory, German is not a regular V2-language. V2 must be suspended in favor of V3 just in the context of these particles/adverbs. The core reason for this step is a semantic one, namely that focus particles must have sentential scope. [Prt NP], [Prt PP] etc. are believed to be incompatible with this semantic requirement because Prt scopes over XP while XP fails to be a proposition.

The account of discourse particles in small PrtPs that was given in section 4 shows how difficulties with scope can be overcome in a more elaborate syntactic theory. The small PrtP is only the visible part of a richer structure in which Prt takes scope in a lower position than where it ultimately appears phonetically. One can be almost sure that a similar account is available for focus particles.[5] In the present context it may suffice to see that just like discourse particles focus particle can also attract an emphatically marked XP to their left. This gives rise to a surface structure by which the particle does not overtly c-command the associated focus. In German, examples abound in which focused XPs raise to the specifier of a focus particle for extra emphasis. The following are from standard newspapers to which I have added relevant focus information by using capitals.

(38) [SELTEN **nur**] hat er Antworten anzubieten
      rarely   only has he answers   to-offer
      "Only rarely can he offer answers"

(39) [EINE SEKUNDE **nur**] hat den monatelangen Kampf mit   dem Deutschen
      one   second   only has the month-long   fight   with the   German
      Tennis-Bund ... zunichte   gemacht
      tennis-league   functionless made
      "Only one second wrecked the month-long fight with the German tennis league"

How the bracketed phrases before the finite verb could be analyzed as non-constituents as suggested by Büring & Hartmann, 2001 is unclear.

This brief detour to focus particles bears on the controversy about the categorial status of discourse particles. If discourse particles were adverbs as argued by Cardinaletti (2007), they would presumably not be able to play the role

---

[5] It would be conceptually odd to have no connections between focus particles and discourse particles as many of them have a life in both domains and are historically closely affiliated with each other.

in the formation of small PrtPs that we have examined in the previous section.[6] The constituency that we have observed there strongly suggests that the discourse particles under consideration are heads. How could they otherwise merge with XPs? Notice that merger with *bona fide* adverbs rarely leads to acceptable results.

(40) *[WEN gestern]   hat Karl getroffen?
      who   yesterday has Karl met
      intended: "Who did Karl meet yesterday"

(41) *[DEN ANTON vermutlich] hat Karl gestern   getroffen?
      the Anton   perhaps   has Karl yesterday met
      intended: "Perhaps Karl met Anton yesterday".

Thus, I consider the data in connection with small PrtP in section 4 as well as the often ignored data in (38) and (39) concerning focus particle as strong evidence for the head status of discourse particles, - at least of those that have been considered so far.[7] The following section will corroborate the analysis that has emerged so far.

---

[6] Cardinaletti is, of course, aware of the properties of discourse particles listed in section 2. She attributes a special status of "weak adverbs" to them. Of course, a theory which can do without such assumptions would have more credibility.

[7] Bare phrase structure (BPS) has formalized a message that has been known for quite a while, namely that categories may occasionally oscillate between heads and phrases. Consider here the German particle *schon* which must be a head in (i) but – due to the V2 constraint – an XP in (ii).

      (i)   [WO **schon**] wird er sein?!
               where SCHON will he be
               „Where will he be after all? (we all know!)"
      (ii)   [**Schon**] hat er das Tor   erreicht
               SCHON has he the gate reached
               "He has already reached the gate"

One response to this apparent non-uniformity could be that head status should nevertheless be seen as rigidly determined. This necessitates the assumption of empty projections. In (ii) there should be an empty temporal XP which is selected by the particle *schon*. The answer of BPS is (or should in my view be) that head status is a consequence of Merge. According to this theory the categorial status of an item is contextually determined. For various reasons which I cannot discuss here, I tend to accept the latter theory. For further discussion cf. Bayer (2002) and Bayer & Brandner (2008).

Josef Bayer

6. Mixed Constituency

Recall that Prt° can be merged with VP/*v*P. This leads to the standard case which we may refer to as the *big PrtP*. A particle that occurs in the context of a wh-question may, however, also be merged with an emphatically accented wh-phrase. This process leads to what we have called a *small PrtP*. The expectation is then that the two processes can combine. The German data verify this expectation.

(42) [Warum **denn**] hätte   er das [**nur**   [~~warum denn~~ sagen sollen]]?
    why   DENN had$_{SUBJ}$ he this NUR   say   should
    "Why on earth should he have said that (I am wondering)?"

Here the small PrtP *warum denn* will activate a PrtP above the big PrtP headed by *nur*. The proper derivation is indicated in (43).

(43) [Warum **denn**] hätte er das [$_{PrtP1}$ ~~warum denn~~ [$_{PrtP2}$ **nur** [~~warum denn~~   sagen sollen]]]?

Nur heads the big PrtP (PrtP2), while the small PrtP headed by *denn* passes through the left edge of a big PrtP (PrtP1) in which *denn* is activated and takes scope. This confirms the attested unique order *denn* < *nur*.

Interestingly, wh-movement may also lead to a linear reversal of the ordering which discourse particles strictly follow. (44) is grammatical although the linear order is now the reverse of the expected order: *nur* < *denn*.

(44) [Warum **nur**] hätte   er das [**denn**   sagen sollen]]?
    why   NUR had$_{SUBJ}$ he this DENN say   should
    "Why on earth should he have said that (I am wondering)?"
                                    (semantically hardly distinguishable from (42)/(43))

The present account accomodates this puzzling phenomenon straightforwardly as a case of reconstruction. In the representation in (45), the small PrtP *warum nur* has left a copy in PrtP2 which is properly ordered below PrtP1:

(45) [Warum **nur**] hätte er das [$_{PrtP1}$ **denn** [$_{PrtP2}$ ~~warum nur~~ [$_{vP}$ ~~warum nur~~ sagen sollen]]]?

7. Conclusion

The present investigation of discourse particles in questions of German has lead to a number of generalizations which may prove to be useful for a deeper understanding of the relation between clause structure, its functional organization

and semantic/pragmatic interpretation. The grammar of discourse particles in German wh-questions has revealed the following features.

◆ Discourse particles seem to be rigorously ordered in a functional skeleton which can be interleaved with topic XPs. Arguably, each particle heads its own particle phrase ("big PrtP").

◆ Next to this option, discourse particle can be merged with emphatically focused wh-phrases ("small PrtP"). This process can be applied recursively, the particles retaining the rigid ordering that has already been observed in the formation of "big PrtPs".

◆ The "small PrtP", which seems to be the result of emphasis marking, maps onto the "big PrtP" via reconstruction (copy movement).

◆ The interrogative force system is complemented by information from the particles via phase-based local agreement which is run in cycles on the basis of wh-movement. Contrary to mainstrean assumptions, discourse particles can thus arise in embedded clauses which do not belong to the force projection but may nevertheless "communicate" with the force projection of the root clause.

◆ Force is not monolithically established in the left periphery but may in principle communicate with projections from unboundedly distant particle projections.

The grammar of discourse particles in German wh-questions shows a rather intricate design in which wh-movement plays a key role. This design appears to be far from universal. In various languages, discourse particles are confined to the matrix clause and may even be banned from non-peripheral positions (cf. Haegeman (2009) for West-Flemisch and Munaro & Poletto (2009) for Italian and Italian dialects). Bangla has agglutinative clause-final or clause-medial particles which appear to be positionally fixed like German discourse particles and have constituents move to their left (cf. Dasgupta (1984; 1987) who suggested the rather appropriate term "anchors" for them). Although nothing of this work could be presented here, it should be clear that the development of a syntactic typology of discourse particles could complement and modify our present understanding of clausal architecture at the syntax-semantics interface.

Josef Bayer

References

Bayer, Josef. 1996. *Directionality and Logical Form. On the Scope of Focusing Particles and Wh-in-situ*. Dordrecht: Kluwer.

Bayer, Josef. 1999. Bound focus in German or how can association with focus be achieved without going semantically astray? In: G. Rebuschi & L. Tuller (eds.) *The Grammar of Focus*. Amsterdam: John Benjamins. 55-82.

Bayer, Josef. 2001. Asymmetry in emphatic topicalization. In: C. Féry & W. Sternefeld (eds.) *Audiatur Vox Sapientiae*. Berlin: Akademie-Verlag. 15-47.

Bayer, Josef. 2002. Minimale Annahmen in Syntax und Morphologie. *Linguistische Arbeitsberichte* 79, Institut für Linguistik, University of Leipzig. 277-297.

Bayer, Josef and Eleonore Brandner. 2008. On Wh-Head-Movement and the Doubly-Filled-Comp Filter. In: C. B. Chang & H. J. Haynie (eds.) *Proceedings of the 26th West Coast Conference on Formal Linguistics*. Somerville: Cascadilla Proceedings Project, 87-95.

Bayer, Josef and Hans-Georg Obenauer (submitted), Discourse Particles, Clause Structure, and Question Types. ms. 2009. University of Konstanz & UMR 7023 CNRS / Univ. Paris 8.

Behaghel, Otto. 1932, *Deutsche Syntax. Wortstellung. Periodenbau*, volume 4. Heidelberg: Carl Winter's Universitätsbuchhandlung.

Büring, Daniel and Katharina Hartmann. 2001. The syntax and semantics of focus-sensitive particles in German. *Natural Language and Linguistic Theory* 19: 229-281.

Cardinaletti, Anna. 2007. Für eine syntaktische Analyse von Modalpartikeln. In: E.-M-Thüne & F. Ortu (eds.). *Gesprochende Sprache – Partikeln*. Frankfurt: Peter Lang. 89-101.

Cinque, Guglielmo. 1999. *Adverbs and Functional Heads: A Crosslinguistic Perspective*, Oxford: Oxford University Press.

Coniglio, Marco. 2005. *Deutsche Modalpartikeln: eine syntaktische Analyse*. MA thesis, Università Ca' Foscari, Venezia.

Dasgupta, Probal. 1984. Bangla emphasizers and anchors. *Indian Linguistics* 45, 102-117.

Dasgupta, Probal. 1987. Sentence particles in Bangla. In: E. Bashir, M. Deshpande and P. Hook (eds.), *Selected Papers from SALA 7*, distributed by Indiana University Linguistics Club.

Gabelentz, Georg von der. 1901. *Die Sprachwissenschaft, ihre Aufgaben und Methoden*. reprint 1969. Tübingen: Narr.

Haegeman, Liliane. 2009. The cartography of discourse markers in West-Flemish. Ms. Department of English, University of Ghent.

Jacobs, Joachim. 1983. *Fokus und Skalen: Zur Syntax und Semantik der Gradpartikeln im Deutschen*, Tübingen: Niemeyer.

Munaro, Nicola & Cecilia Poletto. 2009. Sentential particles and clausal typing in the Veneto dialects. In: P. Cook, W. Frey & C. Maienborn (eds.), *Dislocated Elements in Discourse: Syntactic, Semantic, and Pragmatic Perspectives*. New York: Routledge.

Obenauer, Hans-Georg. 2004. Nonstandard wh-questions and alternative checkers in Pagotto. In: Horst Lohnstein & Susanne Trissler (eds.), *The Syntax and Semantics of the Left Periphery*. Berlin: de Gruyter. 343-383.

Paul, Waltraud. 2008. Sentence Particles from a Chinese perspective. Workshop *Looking for particles*, Venice, 4 February 2008.

Potts, Christopher. 2005. *The Logic of Conventional Implicatures*. Oxford: Oxford University Press.

Rothstein, Susan. 1991. Heads, projections, and category determination. In: K. Leffel and D. Bouchard (eds.), *Views on Phrase Structure*. Dordrecht: Kluwer.

Thurmair, Maria. 1989. *Modalpartikeln und ihre Kombinationen*. Tübingen: Niemeyer.

Variation and the Emergence of Marked Syllable Structure[*]

T. Temsunungsang

The English and Foreign Languages University, Hyderabad

Abstract

In this paper, we examine two aspects of Ao phonology: foot structure and the emergence of complex margins. Firstly, we examine a case of complex margins in Ao where though the syllable canon is maximally CVC, complex margins can emerge as a result of schwa deletion. Such a phenomenon, albeit in free variation, presents new evidence towards the structure of the foot in Ao phonology, which has not been dealt with in earlier literature (Gowda 1978, Coupe 2007; Temsunungsang 2006, 2008). We show that the deletion pattern points to a trochaic foot structure in Ao with a right edge alignment. Secondly, we examine variation as a case of emerging markedness by subscribing to the notion of partial ordering and ordering of subsets of constraints. We will show that variation results due to the non-ranking of faithfulness constraints with respect to two types of markedness constraints. Such unranking of faithfulness constraints result in three different optimal outputs depending on how the faithfulness constraints are ranked over the markedness constraints.

1. Introduction

Free variation in languages has been one of the lesser studied areas in phonological theory. While the phenomenon of free variation has been ascribed to be closely related to factors such as language contact and other social variables like age, sex, gender, etc, the role of grammar needs to be examined in such occurrences. If the grammar is responsible, then how is it to be accounted for? We believe that the variation observed in Ao is a shift in the grammar of the language where rhotic clusters are emerging as a result of deletion in spite of the high ranking constraint on complex margins.

In this paper, we look at Ao[1], a Tibeto-Burman language of Northeast India, where deletion in certain environments leads to consonant clusters not

---

[1] In Ao, there are three main dialects namely Chungli, Mongsen and Changki. In this paper, by Ao, I refer only to the first two dialects.

attested elsewhere in the language. The syllable in both the dialects, Chungli and Mongsen, is maximally CVC in structure. This emergence of marked structures, albeit in free variation, presents new evidence towards the structure of the foot in Ao phonology, which has not been dealt with in earlier literature (Gowda 1978, Coupe 2007; Temsunungsang 2006, 2008). In this paper, we examine two aspects of Ao phonology: foot structure and the emergence of complex margins.

In section 2, we introduce the process of schwa deletion and present arguments for treating it as deletion. Section 3 briefly looks at the foot in Ao. Section 4 examines the minimality conditions for both dialects while section 5 gives an analysis of the rhythmic pattern and foot structure. We present an Optimality Theoretic analysis of variation in section 6 and conclude with section 7.

2. Schwa Deletion

In Ao, schwa deletion is a common phenomenon, [ə] being a variant of [ɯ] and the least marked of the five vowels [i, u, a, ə, ɯ]. The forms below are in free variation and occur in normal speech.

(1) a.   tə́rúk   ~   trúk   'six'[2]

   b.   tárət   ~   tárt   'heavy'

In 1(a), deletion results in a Cr onset cluster while in 1(b), deletion results in an rC coda cluster. The language does not attest complex onsets or codas except for the ones above which is considered a result of deletion. In the above examples, the alternative argument could be that consonant clusters are underlyingly present and undergo simplification through schwa insertion. We present four arguments to suggest that deletion is a better option to postulate rather than insertion.

Firstly, no other clusters are attested in the language except /stop + r/ as complex onset and /r + stop/ as complex coda.

Secondly, as will be discussed in later sections, the site of schwa deletion is predictably the weak position. It would be hard to explain how clusters are underlying only in certain positions.

Thirdly, no other vowel ~ zero alternation is attested (2a) while the schwa is prone to deletion in certain morphological environments like prefixation as in 2(b).

---

[2] In both dialects, three contrastive tones are attested; High, Mid and Low represented by the diacritics ´, ¯ and ` respectively. Tones are omitted where irrelevant.

(2) a. li + aŋ   => liaŋ 'live, imp'   b. tə + asa => tasa   'neg.shout'

Lastly, the schwa is seldom the epenthetic vowel elsewhere. In loanwords for instance, /u/ and /i/ are epenthetic vowels in cluster simplification while in Chungli, /a/ is the vowel which is augmented to satisfy minimality requirements.

(3)   tu.riŋ   'drawing, picture'
       is.kul   'school'
       -səm   =>   a.səm 'wear'

Though such epenthesis is driven in different morphological environments (loanwords, bound verb roots), it sheds light on why clusters are not underlying but a result of deletion.

3. Feet in Ao

Studies on the minimal word in Ao (Sanyal 2005; Temsunungsang 2008) have shown that sub-minimal words are tolerated in Chungli nouns and Mongsen verbs.

(4) Chungli: nuk   'knife'   ku   'hair'
      Mongsen: nak   'guard'   li   'live'

However, a minimal disyllabic requirement is attested for verbs in Chungli and nouns in Mongsen.

(5) Chungli: anak   'guard'   aluk   'wash'
      Mongsen: anuk   'knife'   məzət 'anger'

The structure of foot and its rhythmic pattern is one aspect of Ao phonology which has been neglected in previous studies (Gowda 1978, Coupe 2007; Temsunungsang 2006, 2008). This is perhaps due to the lack of evidence with respect to stress; intensity, duration, pitch. Unlike other tonal languages (de Lacy 1999), tones in Ao do not show any evidence of stress being attracted to certain positions in a word. So, is the notion of foot irrelevant in the language?

        As Yip (1994) suggests, a single phenomenon is sufficient to postulate such notions as foot, syllable and mora in a grammar. According to Yip (1994), many of the Chinese dialects show little traditional evidence for prosodic categories below the phonological phrase. She argues that the notions of mora, syllable and

foot, however, play an important role in these dialects, concluding that these units must be part of the inventory supplied by universal grammar.

Considering the lack of traditional evidence in Ao for stress, we must look at other evidence, if any. Firstly, the notion of foot is strongly evident from the strict adherence to a disyllabic minimum in Chungli verbs and Mongsen nouns (section 4) Secondly, deletion points to specific rhythmic patterns (section 5).

4. Minimal Words

It is a generally accepted view that lexical categories across languages meet certain requirements with regard to their minimal size/length. This assumption is arrived at from the Prosodic Hierarchy and complemented by the Foot Binarity Principle. According to the Prosodic Hierarchy (Selkirk 1980, 1984; Nespor and Vogel 1986; McCarthy and Prince 1986) prosodic constituents are organized in a hierarchical order.

(6)   Pr WD   Prosodic Word
              |
            FT   Foot
              |
             σ   Syllable
              |
             μ   Mora

Though there has been considerable agreement on the Prosodic Hierarchy, its universal status has been a subject of much debate. While the presence of all the levels in a language is argued for by Nespor and Vogel (1986), many scholars starting from Selkirk (1980) opine that all levels may not be utilized in a language. (cf. Hyman (1982), Auer (1994), Kleinhenz (1996)).

The Prosodic Hierarchy conforms to the Strict Layering Hypothesis where a prosodic constituent of level i dominates a lower level i-1 such that all prosodic constituents are exhaustively parsed. A principle that conforms to this hypothesis is the Foot Binarity principle (McCarthy and Prince 1991, Prince 1980).

Foot Binarity: 'Feet must be binary under syllabic or moraic analysis'

This Principle dictates that a foot must consist of a heavy syllable or two light syllables in quantity sensitive languages while quantity insensitive or syllabic languages must consist of two syllables. In other words, a foot must be binary branched. Thus, if a Prosodic word contains a foot and a foot consists of two syllables or two moras, then by transitivity, a prosodic word is bimoraic or

disyllabic. Irrespective of whether a grammar parametricizes for strong or weak layering in the Prosodic Hierarchy, it is generally assumed that foot binarity is rarely called into question.

In a brief section on Chungli and Mongsen underived words below, we will see that while Chungli verbs and Mongsen nouns conform to the universal minimality conditions, Chungli nouns and Mongsen verbs do not adhere to the Foot Binarity Principle. In other words, minimality conditions are violable depending on the grammatical category allowing for degenerate feet to surface in the language. To our knowledge, such a case where minimality requirements differ depending on the grammatical category has not been attested elsewhere in the literature.

4.1 Chungli

In Chungli, a major evidence for the existence of foot structure comes from vowel augmentation in bound verb roots (see Temsunungsang and Sanyal 2005 who argue for truncation). These roots must undergo vowel augmentation resulting in disyllabic forms.

(7) -mət => amət   'boil'   -mən => amən 'sit'   -wuk=>awuk 'sweep'

Hence, while bound roots form degenerate feet, augmentation helps in building binary feet.

Nouns in Chungli on the other hand are mostly monosyllabic. Many of these words are of CV or CVC$_{stop}$ structure where the assumption is that the coda does not contribute to weight and hence form subminimal words.

(8)   nùk 'knife'   sə̀n 'money'   mì 'fire'

What we see in Chungli nouns are degenerate feet or subminimal prosodic words; a highly ranked universal constraint but violable in Chungli.

4.2 Mongsen

In Mongsen, we see a similar pattern though it is an exact mirror image of what we find in Chungli. While Mongsen verbs follow the pattern of Chungli nouns (9), Mongsen nouns follow the pattern of Chungli verbs.

(9)   tə̄n 'sing'   cə̀p 'cry'   sā 'say'

Here again, the assumption is that coda stops do not carry weight and hence, Mongsen verbs allow for subminimal words.

In most nouns, an initial /a/ is attested which has been treated as a non-relational prefix (Coupe 2007). The initial vowel productively deletes in compound and possessive formations.

(10)   an 'hen' + asaʔ 'meat'=>   an-saʔ 'chicken meat'

    niŋ 'your' + alu 'field'   =>   niŋ-lu 'your field'

Hence, just as in Chungli, nouns in Mongsen can be considered to be bound roots underlyingly /-saʔ/ which undergo /a/ augmentation in citation forms to satisfy prosodic requirements.[3]

5. Rhythmic Pattern

The phenomenon of deletion resulting in complex margins is one evidence which suggests that the rhythmic pattern of Ao is a Trochee. In a disyllabic word, the deletion is attested in the initial as well as final syllable, *i.e.,* wherever it occurs as in (1). However, consider (11) where there are two schwas within a word.

(11)   a.   pə́rət ~   pə̂rt   *prət 'spin'

    b.   pə̀rə̄m ~   pə̀r̄m   *prəm 'jewel'

Given a choice, it is the schwa on the right which is deleted consistently. The tone is however retained on the sonorant /r/. Our hypothesis is that the choice of deletion is driven by rhythmic factors: if there is more than one vowel to be deleted, the one on the weak syllable is deleted. Though our understanding of foot structure in Ao is limited, we believe that the deletion in the second syllable is due to its weak position, i.e., Ao builds a trochaic foot. However, not all sequences of Cər results in cluster formation as will be argued below.

5.1 Prefixation and Foot Formation

As seen above, onset clusters can consist of stop + r. While the stops /p/ and /k/ have limited occurrence, the occurrence of /t/ is prevalent due to the presence of the homophonous prefix form /tə̀-/ for negation, nominaliser and the relational prefix for body parts. In other words, the prefixation of these forms to /r/ initial

---

[3] In both Chungli verbs and Mongsen nouns, monosyllabic forms are attested which are however considered to be a result of contraction and hence underlyingly disyllabic (Temsunungsang, 2009).

stems must trigger deletion, resulting in clusters. We examine only the negative prefix in this paper.

The negative prefix which has a low tone, behaves in different ways. In Chungli, verbs are divided into four different types based on their structure and morphological behavior as is shown below (Temsunungsang 2006, Sanyal 2005):
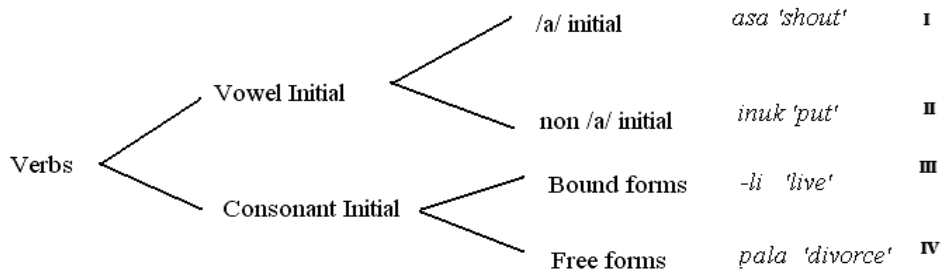
On prefixation to /a/ initial verbs (type I), the schwa always deletes, at times along with the tone as in (12).

(12) a. tə̀ + ásà   =>   tásà   'neg.shout'

   b.   tə̀ + àsə̄m   =>   tàsə̄m 'neg.run'

The same is observed with other vowel initial stems (type II and III), with the additional option of the stem vowel being copied and retaining the prefixal L tone.

(13)   tə̀ +ī̄   =>   tə̄ī ~ tĭ̄ī   (type III)

    tə̀ + ínùk   =>   tə́ínùk ~ tĭ̄ínùk (type II)

Hence, the condition for schwa deletion is quite clear. Firstly, deletion occurs only when it is in the environment of the rhotic.

Contrary to the above (12 and 13), the schwa is not deleted in certain environments. For instance, prefixation to bound roots (type III) seldom result in schwa deletion.

(14)   -rūng   tə̀rūŋ   *truŋ   'neg.burn'

    -ràk   tə̀ràk   *trak   'neg.pull out'

Differing from the behavior in bound roots (14), prefixation to free forms (type IV) can result in schwa deletion and the low tone of the prefix is not lost but shifted to the nearest vowel as in (15).

(15)   rárà   tə̀rárà ~ trǎrà   'neg.fight'

        rə̄nə̄m   tə̀rnə̄m ~ trə̀nə̄m   'neg.ready'

This difference in behavior of the prefix (cf. 14 and 15) depending on the root can be explained by minimality conditions imposed on Chungli verbs and the reorganization of the foot as a result of prefixation. Minimally, verbs are disyllabic and hence the initial syllable /tə/ is preserved in (14) above to avoid degenerate feet. In (15), this minimality requirement stands fulfilled within the root and hence the option of schwa deletion is exercised. On the other hand, Mongsen verbs having no minimality requirements, schwa deletions are attested more frequently than in Chungli making the above forms in (14) possible[4].

We examine two possible ways in which the prefix + root can be schematized and argue for right edge alignment of foot below.

(16) a. (tərə)nəm   =>   trə.nəm/tər.nəm

     b. tə (rənəm)   =>   trə.nəm/tər.nəm

If we build the foot from the left edge as in (16a), we need to explain the deletion of [ə] from a strong position for [trə.nəm] while [tər.nəm] does not pose a problem since deletion will be in the weak position.

If the foot is built from the right edge as in (16b), [trə.nəm] can be considered to be a case of schwa deletion and adjunction to the foot. [tər.nəm], on the other hand, must be explained as a case of deletion in the strong position.

While both cases seem problematic, we choose the second explanation where the schwa in [tər.nəm] will be considered to be a case of vowel metathesis to avoid complex clusters. Hence,

(17)   tə̀ (rə̄nə̄m) => (trə̀.nə̄m)   adjunction to foot

                         ⇓

                      (tə̀r.nə̄m)   metathesis

---

[4] In Mongsen nouns, not many words of Cər sequence in disyllabic words are attested to satisfactorily test whether it avoids cluster formation to satisfy minimality conditions. For instance in [kəraʔ] 'tortoise', no deletion is attested, while in [tʰəra] 'ten' the deletion is consistent.

While metathesis is not a commonly attested, cases are also attested in Mongsen. For eg, tə-runəm 'adj.aroma' can surface as [tur.nəm] or [tru.nəm].

The right edge salience is also borne out in other aspects of the phonology. Firstly, in bitonal and tritonal stems, the lexical tone is at the right edge of the word. The left edge gets its tonal specification through spreading.

(18)   -mə̄n   =>   āmə̄n   'ripe'   (Chungli)

        -núk   =>   ánúk   'knife'   (Mongsen)

Secondly, in a tritonal stem, the insertion of a default M tone is always at the initial (left edge).

(19)   àrə̄kū 'church'   (both dialects)

        āpúsù 'start'   (Chungli)

Lastly, suffixes which occur at the right edge seem to dictate the tonal pattern of the word in derived forms.

(20)   -zə̀k + āŋ   =>   zə̄kāŋ   'beat'   (Chungli)

        yàk + āŋ   =>   yākāŋ   'beat'   (Mongsen)

6. Variation in Optimality Theory

In Optimality Theory (OT) (Prince and Smolensky 1993), three components of the grammar interact to give the optimal output: (1) the lexicon which supplies the input to (2) the generator (GEN), generating an infinite number of output candidates. The candidates are then fed into the (3) evaluator (EVAL) which evaluates the candidates against a set of ranked constraints. In principle, EVAL must choose only one candidate as the optimal output based on the ranking of constraints. How then must free variation be accounted for by a grammar of ranked constraints. While Prince and Smolensky (1993) agree on the possibility of crucial non-ranking in languages, it is dismissed based on the lack of evidence. This hypothesis is however not shared by researchers who claim that nonranking is crucial for variation in languages.

There have been a number of positions in OT on the nature of (non)ranking to account for free variation. Some of them include Kiparsky (1993), perhaps the earliest work on variation in OT, who assumes that variation comes from

competing grammatical systems of individuals or community, postulating partial ranking of constraints in a hierarchy. Hence for English t/d deletion, three different rankings predict three categorical deletion patterns. Hammond (1994), based on Idsardi's (1992) argument that an exclusively-constraint based grammar is problematic since multiple outputs can be generated from an incomplete constraint hierarchy, argues how variation is a result of multiple outputs in an incomplete constraint hierarchy for variable stress in Walmatjari; Antilla (1997) makes use of partial ordering for Finnish stress while Reynolds (1994); Nagy and Reynolds (1997) argue for floating constraints where a constraint is unranked with respect to a set of constraints. (see Boersmo 1998, Boersmo and Hayes 2001, Antilla 1997, Antilla and Cho 1998 for more constrained models).

In the model proposed by Reynolds (1994); Nagy and Reynolds (1997), a floating constraint is unranked with respect to a subset of constraints. Hence, if constraint D is unranked with respect to A>>B>>C, four possible rankings emerge allowing for four possible outputs.

A>>B>>C>>D, A>>B>>D>>C, A>>D>>B>>C, D>>A>>B>>C

A similar proposal is extended in Antilla (1997), Antilla and Cho (1998) with a slight variation, where only adjacent constraints can be unranked. Hence, if constraint D was unranked, it can be so only with respect to one of the constraints A, B or C.

A>>B>>C>>D,   A>>B>>D>>C

An important aspect of this proposal is that subsets of constraints are ranked with respect to each other but constraints within subsets are not.

set 1 >> set 2 >> set 3 (A, B, C)

Hence, if there are three subsets of constraints set 1, set 2 and set 3, and set 3 has three constraints, A, B and C, the nonranking of these three constraints in set 3 allows for variation. This approach allows for a number of rankings depending on the number of constraints in a subset.

In this analysis, we follow the notion of partial ordering and ordering of subsets of constraints in a more constrained manner. While the grammar may have subsets of constraints, we argue that only certain constraints need to be unranked constraining the number of ranking within this subset. If A, B, C and D are constraints within a subset and are unranked, there are 24 possible rankings available leading to 24 possible outputs. However, if A>>B and C, D are unranked with respect to A and B, the possible rankings are reduced to twelve.

A>>B>>D>>C	D>>A>>B>>C	C>>D>>A>>B
A>>B>>C>>D	D>>A>>C>>B	C>>A>>D>>B
A>>D>>B>>C	D>>C>>A>>B	C>>A>>B>>D
A>>C>>B>>D	A>>C>>D>>B	A>>D>>C>>B

We will show that in Ao, A and B are markedness constraits which are strictly ranked with respect to each other while C and D are faithfulness constraints which are free. While the twelve outputs are a possibility given the ranking above, only three possible outputs surface due to violation of higher ranked constraint subsets in the hierarchy.

In relation to the emergence of marked structure, it is necessary to make note of the ranking logic in Benua's (1997) The Emergence of the Relatively Unmarked (TETRU) and McCarthy and Prince's (1994) The Emergence of the Unmarked (TETU) in reduplication.

TETRU: OO-IDENTITY >> M >> IO-FAITH
TETU: IO-faithfulness >> Well-formedness >> BR-identity

What is common in the two phenomena is that markedness dominates one kind of faithfulness which is in turn dominated by some faithfulness constraint.

While OO-identity or BR-identity is irrelevant in Ao, we will show the emergence of marked structures is a consequence of not ranking faithfulness constraints over two types of markedness constraints. This results in three possible variants depending on where faithfulness is ranked.

6.1 Emergence of Complex Margins

In the section below, we spell out the constraints necessary to account for the variation and emergence of marked structures in certain environments. In the previous sections, we have looked at one case of variation occurring as a result of schwa deletion in specific environments of stop-schwa-rhotic sequences and rhotic-schwa-stop sequences. Note that only the schwa is prone to deletion, other vowels remaining unaffected. This requires the division of the faithfulness constraint MAX into more specific constraints: MAX-IO{i. u, a} and MAX-IO{ə} with the ranking order;

(21)   MAX-IO{i. u, a} >> MAX-IO{ə}
    MAX-IO{i. u, a}   No deletion of /i/, /u/ and /a/
    MAX-IO{ə}   No deletion of [ə]

In addition, given a choice, deletion always takes place in a non-head position as seen in section 5. This is accounted by HEADMAX which disallows deletion in heads and must be ranked with respect to MAX-IO{ə}.

(22)   HEADMAX[5]   Every segment in the prosodic head must have a correspondent in the output

tərət 'bone'	HEADMAX	MAX-IO{ə}
a. trət	*	*
b. ☞ tərt		*

Tableau 1

In tableau 1, candidate (a) violates HEADMAX which has a deletion in the head position while candidate (b) violates only the lower ranked constraint MAX-IO{ə}. In fact, the ranking does not matter here as Max-IO{ə} >> HEADMAX will still give the desired results. However, we have not considered a third candidate, [tərət] which is most faithful to the input and a possible free variant form. This candidate emerges the winner if evaluated. Hence, there must be some other constraint which interacts with the above faithfulness constraints. We first look at two markedness constraints *STRUC and *COMPLEX below and returning to the above in this section.

We have seen that consonant clusters do not occur underlyingly, suggesting *COMPLEX to be a high ranked constraint. It is only when deletion occurs in certain environments that clusters emerge. We postulate that schwa deletion is motivated by the grammar's need to reduce phonological structure wherever the environment is fulfilled. The constraint *STRUC thus interacts with *COMPLEX where *STRUC >> *COMPLEX.

---

[5] As pointed out by Paul Kiparsky, HeadMax assumes that the input comes with a prosodic head (or foot structure), which can achieved by serial based OT. The question of what the input consists of, whether in serial or parallel based OT, needs further analysis. However, due to space constraints, we do not deal with this issue in this paper. One can refer to Temsunungsang (2009).

(23)   *STRUC   No phonological structure
         *COMPLEX   No complex margins

tәruk 'six'	*STRUC	*COMPLEX
a. tәruk	*!	
b. ☞ truk		*

Tableau 2

In the above tableau 2, candidate (a) violates *STRUC resulting in candidate (b) as the optimal output. Other possible candidates, [tәrk] and [turk] would violate the higher ranked constraint MAX-IO{i. u, a} and DEP-IO{i.u.a}[6] respectively and hence are not shown in the tableau. Going by the tableau above, MAX-IO{ə} must be ranked either below *COMPLEX or *STRUC. Ranking it above the two markedness constraints will give the wrong output, *i.e.*, candidate (a)

(24)   MAX-IO{i. u, a} >> *STRUC >> (Max-IO{ə}) >> *COMPLEX >> (Max-IO{ə})

However, candidate (a) is also a possible free variant which can be achieved by a reversal of ranking, *i.e.*, *COMPLEX >> *STRUC. However, rather than reverse the ranking order, we have the option of allowing MAX-IO{ə} to be unranked with respect to *STRUC and *COMPLEX. This turns out to be the better explanation since the failure of deletion results in a more faithful output [tәruk].

(25) Preliminary Ranking

MAX-IO{i. u, a}>>(MAX-IO{ə}) >> *STRUC >> (MAX-IO{ə}) >> *COMPLEX >> (MAX-IO{ə})

By this ranking logic, if MAX-IO{ə} is ranked above the markedness constraints, the most faithful output [tәruk] wins over candidates with clusters. If MAX-IO{ə} is ranked below any one of the markedness constraints, outputs with clusters [truk] win over the most faithful output.

---

[6] DEP-IO{i,u,a} violates /a/ augmentation in verbs. However, verbs satisfy a higher ranked constraint: FT-BIN(verb).

The other constraint which needs to be ranked with respect to the above constraints is HEADMAX. We have shown in (21) where MAX-IO{i. u, a} >> MAX-IO{ə}. With respect to HEADMAX and MAX-IO{ə}, the ranking was not crucial as seen in tableau 1. With the assumption that MAX-IO{i. u, a} is higher ranked, we will show below how the non-ranking of HeadMax and Max-IO {ə}is crucial for the variation patterns attested in Ao.

Note that there are three positions in which MAX-IO{ə}can possibly occur as seen in our preliminary ranking. This would also give HEADMAX three positions in which to occur in the constraint hierarchy. We test this out with [tə-rənəm] 'neg.ready' in the tableau below where the four constraints will be considered as one constraint subset:

Case 1:　　　HEADMAX >> MAX-IO{ə} >> *STRUC >> *COMPLEX

tə-rənəm 'neg-ready'	HEADMAX	MAX-IO{ə}	*STRUC	*COMPLEX
a. ☞tərənəm			*	
b. trənəm		*!		*
c. tərnəm	*!	*		

Tableau 3

In tableau 3, all the three candidates are possible variants; candidate (c) with schwa deletion in the head position violates HEADMAX, candidate (b) has one violation for schwa deletion MAX-IO while candidate (a) emerges the most optimal with a violation of lower ranked *STRUC. The ranking of HEADMAX is not crucial here and shown in brackets. (HEADMAX) >> MAX-IO{ə} >> *STRUC >> *COMPLEX

Case 2:　　　*STRUC >> HEADMAX >> MAX-IO{ə} >> *COMPLEX

tə-rənəm 'neg-ready'	*STRUC	HEADMAX	MAX-IO{ə}	*COMPLEX
a. tərənəm	*!			
b. ☞trənəm			*	*
c. tərnəm		*!	*	

Tableau 4

In case 2, candidate (b) emerges as the optimal output with candidate (a) and (c) violating higher ranked constraints *STRUC and HEADMAX. Here too, the ranking of HEADMAX above or below *STRUC is inconsequential since irrespective of the ranking, the optimal output will always be candidate (b). (HEADMAX) >> *STRUC >> (HEADMAX) >> MAX-IO{ə} >> *COMPLEX >> (HEADMAX)

Case 3:   *STRUC >> *COMPLEX >> HEADMAX >> MAX-IO{ə}

tə-rənəm 'neg-ready'	*STRUC	*COMPLEX	HEADMAX	MAX-IO {ə}
a. tərənəm	*!			
b. trənəm		*!		*
c. ☞ tərnəm			*	*

Tableau 5

In case 3, while candidates (a) and (b) fatally violate the markedness constraints, candidate (c) emerges the winner with violation of low ranked faithfulness constraints. Unlike in case 1 and 2, the ranking of HEADMAX is crucial for the optimal out. MAX-IO{ə} can have two positions; above *COMPLEX or below.
*STRUC >> (MAX-IO{ə}) >> *COMPLEX >> HEADMAX >> (MAX-IO{ə})

The above cases further reinforces our postulation that MAX-IO and HEADMAX is unranked in the hierarchy with respect to the markedness constraints yielding three possible outputs. Doing away with the constraint which are not crucial, we have:

(26) MAX-IO{ə} >> *STRUC >> *COMPLEX   =>   tə.rə.nəm

(27) *STRUC >> MAX-IO{ə}>>*COMPLEX   =>   trə.nəm

(28) *STRUC >> *COMPLEX >> HEADMAX   =>   tər.nəm

The ranking of faithfulness over markedness results in the most faithful candidate as in (26). The ranking of Faithfulness between the two markedness

constraints results in emergence of complex margins as in (27) and lastly, the ranking of markedness over faithfulness results in a candidate without clusters but reduced phonological structure as in (28). While these four constraints form a subset of constraints, higher ranked constraints like MAX-IO{i, u, a} form a higher constraint subset with other higher ranked constraints like FT-BIN.

6.2 *COMPLEX: Segments or Clusters

Before concluding, we make a brief note on whether complex clusters can be treated as complex segments. In the analysis of complex onsets and codas, we have seen examples where a complex onset consists of stop + /r/ while a complex coda can consist of /r/ + stop. Most of the examples with complex onsets attest /tr/ while codas have a larger inventory of clusters /rt/, /rk/, /rʔ/, /rp/. In fact, only two words /praŋ/ 'shine' and /pruk/ 'scatter' (in both dialects) are attested where a stop other than /t/ is part of the cluster. While the status of complex codas being clusters is clear, a question arises as to whether /tr/ needs to be treated as a complex segment based on the fact that other clusters are hardly attested in the onset.

A look at the Tibeto-Burman languages show the presence of complex segments in many languages like Mao (Giridhar 1994), Angami (Giridhar 1980) and Lotha (Acharya 1983) attesting complex segments like /pf/ and /bv/, Mizo (Fanai 1992) attesting tl, tr, thl, thr, Dzongkha (Dorjee p.c) attesting /tr, tʰr and Yimchungru attesting /tr/ and almost all languages of the area attesting all or some of the complex segments /ts/, /dz/, /tʃ/ and /dʒ/. Given the prevalence of such complex segments, it is perhaps the case that Ao is a case of complex segments emerging in the onset while complex clusters are emerging in the coda. In fact, the Ungma dialect of Chungli makes a distinction between /tr/ and /t/.

(29)   trɔʔ   'six'   tərɔʔ   'throat'

If this is the case, there seems to be evidence to speculate a further shift in grammar towards complex segments in the onset from clusters. This is perhaps more evident in younger speakers where forms such as [təruk] 'six' are barely in existence. However, we leave the present speculation for later research.

7. Conclusion

In this paper, we have looked at two aspects of Ao phonology: foot structure and variation in syllable structure resulting in complex margins. With respect to foot structure, we postulate that Ao allows for a trochaic system with a right edge

salience. The adjunction of the negative prefix to certain stems creates the environment for variation to occur, resulting in complex margins.

Variation suggests a shift in the grammar of the language by allowing complex margins on the surface and supports the hypothesis that variation is a result of competing grammatical systems (Kiparsky 1993). We postulate that such emergence of complexity is driven by the grammar to reduce phonological structure, under the assumption that Ao must have been a monosyllabic language at one point of time just like many of the Chinese dialects and Southeast Asian languages. Hence, the constraint for reduction of phonological structure *STRUC has remained active in the language.

The influence of *STRUC can also be seen elsewhere. For instance, of the six possible coda consonants, back consonants /k, ŋ/ occur after back vowels /a, u/ while front consonants /p, t, m, n/ occur after [i, ə], presenting a harmonic relation in the rime in terms of [±back]. In addition, an OCP constraint bans labials to co-occur in the onset and coda. But sequences which are otherwise banned, are attested as a result of contraction in (30), again pointing to the dominance of *STRUC in the language.

(30) a. míjə́m =>   míḿ   'love' (contraction)

　　b. āhə̄n =>   āń̄   'hen'   (dialectal difference)

While, we do not go into the details of such forms of markedness in this paper, it is worthwhile to note that syllable contraction is a prevalent process in many of the monosyllabic Chinese languages (Chung 1992), again pointing to our hypothesis of earlier Ao dialects having a strong preference for monosyllabic morphemes. In fact, Ao attests a number of monosyllabic roots and stems in the language as well as the presence of disyllables with prefix like forms as the initial syllable.

In addition, if we look at the descriptive grammars of languages in Northeast India, Ao is perhaps one of the very few languages which do not have clusters. In fact, many languages (see section 6.2) like Mizo (Lalrindiki) and Yimchungru (field notes) attest complex segments like tr, tl etc. Hence, this would suggest that clusters or complex segments are areal features which perhaps have had an influence on Ao.

With respect to variation in OT, we follow the model put forward in Antilla (1997) and Nagy and Reynolds (1997) that constraints within a subset are not ranked. While this is true in Ao, we depart from these models by claiming that markedness constraints need not be unranked, allowing only the faithfulness constraints to be unranked with respect to the markedness constraints. Such a proposal constrains the number of possible outputs which is shown by an analysis of the variation pattern attested in Ao. The three possible variants/outputs are explained by the following ranking logic.

(31)   Faith >> Markedness
        Markedness >> Faith
        Markedness >> Faith >> Markedness

When Faith >> markedness, the output is the most faithful as in (26). When Markedness >> Faith, a less faithful candidate than the most faithful one arises as in (28). When Markedness >> Faith >> Markedness, complex margins arise as in (27).

While one could question the exclusive mobility of faithfulness constraints in a hierarchy, this is perhaps reflected in the well established notions of markedness dominating faithfulness in acquisition. When the literature talks of faithfulness moving over markedness, the assumption is that faithfulness is capable of movement.

References

Acharya, K.P. 1983. *Lotha Grammar*. Mysore: Central Institute of Indian Studies

Auer, P. 1994. Einige Argumente gegen die Silbe als universale prosodische Hauptkategorie. In *Universale phonologische Strukturen und Prozesse,* eds. K. H. Ramers, H. Vater & H. Wode, 55-78. Tübingen: Niemeyer.

Anttila, Arto. 1997. Deriving variation from grammar. In *Variation, Change and Phonological Theory,* eds. Frans Hinskens, Roeland van Hout and Leo Wetzels, 35-68. Amsterdam / Philadelphia: John Benjamins.

Anttila, Arto and Young-mee Yu Cho. 1998. Variation and change in Optimality Theory. *Lingua* 104, 31-56, Special issue on Conflicting Constraints.

Benua, Laura. 1997. Transderivational Identity: Phonological relations between words. PhD dissertation. University of Massachusetts, Amherst.

Boersma, Paul. 1998. Functional phonology: Formalizing the interactions between articulatory and perceptual drives. Ph.D. dissertation, University of Amsterdam, Holland Academic Graphics, The Hague.

Boersma, Paul, and Bruce Hayes. 2001. Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry* 32, 45-86.

Chung, Raung-Fu. 1992. Syllable Contraction in Chinese. Ms, National Kaohsiung Normal University, Taiwan.

Coupe, Alexander Robertson. 2007. *A Grammar of Mongsen Ao*. Berlin: Mouton de Gruyter.

de Lacy, Paul. 1999. Tone and prominence. ROA 333.

Giridhar, P.P. 1980. *Angami Grammar.* Mysore: Central Institute of Indian Studies.

Giridhar, P.P. 1994. *Mao Naga Grammar.* Mysore: Central Institute of Indian Studies.

Gowda, K.S. Gurubasave. 1978. *Ao Grammar*. Mysore: Central Institute of Indian Languages.

Hammond, Michael. 1994. An OT account of variability in Walmatjari stress. Ms., University of Arizona.

Hyman, L. M. 1982. Are there syllables in Gokana?. In *Current approaches to African linguistics. Volume 2,* eds. J. Kaye, H. Koopman, D. Sportiche & A. Dugas, 171-179. Dordrecht: Foris.

Idsardi, W. J. 1992. The Computation of Prosody, MIT Ph.D. dissertation.

Kiparsky, Paul. 1993. An OT Perspective on phonological variation. Handout from *Rutgers Optimality Workshop* 1993, also presented at NWAVE 1994, Stanford University.

Kleinhenz, U. 1996. Zur Typologie phonologischer Domänen. In *Deutsch – Typologisch*, eds. E. Lang & G. Zifonun. Berlin: de Gruyter.

Fanai, Lalrindiki T. 1992. *Some Aspects of the Lexical Phonology of Mizo and English: An Autosegmental Approach*. PhD dissertation, CIEFL: Hyderabad, India.

McCarthy, J.J. and Alan Prince. 1986. *Prosodic Morphology*. Ms:University of Massachusetts.

T. Temsunungsang

McCarthy, J.J. and Alan Prince. 1991 Prosodic Minimality, lecture presented at Univ. of Illinios Conference *The Organization of Phonology*.

McCarthy, J.J and Alan Prince. 1994. The emergence of the unmarked: Optimality in prosodic morphology. In *proceedings of the NELS* 24, 333-379. Amherst: GLSA.

Prince, Alan. 1980. A Metrical Theory for Estonian Quantity. *Linguistic Inquiry* 11, 511-562.

Prince, Alan and Paul Smolensky. 1993. *Optimality Theory: Constraint Interaction Generative grammar.* Rutgers University, New Brunswick, and University of Colorado, Boulder.

Nespor, Marina and Irene Vogel. 1986. *Prosodic Phonology.* Dordrecht: Foris Publications.

Reynolds, William T. 1994. Variation and Phonological Theory. Ph.D. dissertation, Department of Linguistics, University of Pennsylvania.

Sanyal, Paroma. 2005. *Some aspects of the Morpho-Phonology of Tones in Ao: an Optimality Theoretic Approach*. M.Phil Dissertation. CIEFL, Hyderabad.

Selkirk, E.O. 1980. Prosodic Domains in Phonology: Sanskrit revisited. In *Juncture*, eds. M.Aronoff and M.-L. Kean. Saratoga, Calif.:Anma Libri.

Selkirk, E.O. 1984. On the Major Class Features and Syllable Theory. In *Language Sound Structure,* eds. M. Aronoff, R. Oehrle, B. Wilker, and F. Kelly. Cambridge, Mass.: The MIT Press.

Temsunungsang, T. 2006. *Effects of the prefix on the stem.* 28th AICL, 2nd-4th November, 2006. Banaras Hindu University. Varanasi.

Temsunungsang, T. 2008. Tonality and the analysis of sub-minimal words in Ao. In *North East Indian Linguistics*, eds. S. Morey and M. Post. New Delhi, Foundation/Cambridge University Press India.

Temsunungsang, T. 2009. Aspects of the Prosodic Phonology of Ao: An Inter-dialectal Study. PhD dissertation. The English and Foreign Languages University. Hyderabad.

Temsunungsang, T and Paroma Sanyal. 2005. Truncation in Ao Verbal Morphology. *CIEFL Occasional Papers* 11:49-71.

Yip, Moira. 1994. Isolated uses of prosodic categories. In Perspectives in Phonology, eds. Jennifer Cole and Charles Kisseberth, 293-311. CSLI lecture notes, no 51.

Movement Paradoxes and Topic Phrases

Kwang-sup Kim

Hankuk University of Foreign Studies, Republic of Korea

Abstract

There are many instances in which movement turns an ungrammatical sentence into a grammatical one and vice versa. This phenomenon is often referred to as movement paradoxes. However, this paper claims that movement paradoxes are not paradoxes but they follow from the fact that Topic Phrase is a DP/NP, regardless of its original grammatical category, and it must undergo topicalization, since it has a strong topic feature. The constituent with a strong topic feature must not be pronounced, and as a consequence, only the copy at the topic position can be pronounced, which gives rise to all the four types of movement paradoxes.

1. Introduction

Non-DP/NP complements, in contrast to DP/NP complements, display peculiar patterns with regard to movement. Sentence (1a) is ungrammatical because it does not satisfy the selectional restriction that *thinks of* takes a DP/NP-complement only. Surprisingly, however, sentence (1b) is grammatical, even though it appears to violate the same requirement. This is not an idiosyncratic property of the predicate *think of*, but there are many cases in which movement of non-DP/NP complements turns an ungrammatical sentence into a grammatical one and vice versa. The phenomenon, which is often called movement paradoxes, can be categorized into four types.

Type 1
(1) a. *He didn't think of [that he might be wrong].
     b. [That he might be wrong], he didn't think of.
(2) a. *I expect [on this wall] to be hung a portrait of our founder.
     b. [On this wall] I expect to be hung a portrait of our founder.
        (Bresnan (1994b: 98b, 101a))

Type 2
(3) a. *This theory captures [that languages are learnable].
     b. [That languages are learnable] $_i$ is captured t$_i$ by this theory.

(4)   a. *This formulation of the rule expresses/captures/reflects/brings out that
          the nouns behave differently.
      b. That the nouns behave differently is expressed/captured/reflected/brought
          out by this formulation of the rule. (Arlenga 2005: 184)

Type 3
(5)   a. He didn't think [that he might be wrong].
      b.*[That he might be wrong], he didn't think.
(6)   a. It seems that the Giants lost the World Series.
      b. *That the Giants lost the World Series, it seems. (Arlenga 2005: 196)

Type 4
(7)   a. Theodore held [that the chimp was immortal].
      b. *[That the chimp was immortal] was held by Theodore. (Postal 2004: 280)

The sentences in (1-2) are instances of type 1: that is, type 1 is characterized by
the fact that a well-formed string is generated when the mismatched complement
undergoes A'-movement. Type 2 is a case in which a well-formed sentence is
made if the mismatched complement undergoes A-movement. Type 3 is a
reverse of type 1 in the sense that the well-matched complement cannot be
topicalized, and type 4 is a reverse of type 2 in that the well-matched
complement cannot undergo A-movement. One might argue that the examples in
(1-4) may turn out to be not paradoxes, if the complement of each predicate is
not the non-DP/NP constituent at a topic position but a nominal empty category.
After pointing out that the approaches making use of an empty category are
empirically inadequate, this paper explores the possibility of explaining them
while maintaining that the dislocated constituents and their gap are related via
movement.

2.   Koster-Alrenga's Pro-Approach

Koster (1978) takes the view that no movement takes place in the cases
illustrated above. He proposes that the sentential topic phrase is base-generated at
the outer SPEC-C and it is associated with a pro, which is generated in the
thematic position and undergoes A'-movement to the inner SPEC of C.

(8)   [$_{CP}$ topic$_i$ [$_{CP}$ pro $_i$ [$_{TP}$ …verb t $_i$]]]

The pro-approach can straightforwardly deal with the type 1 mismatch. Sentence
(1a) is ruled out because *think of* subcategorizes for a DP/NP only. On the other
hand, the subcategorization of *think of* is satisfied in (1b) if we assume with
Koster (1978) that the sentence is analyzed as (9), in which pro satisfies the

requirement that *think of* take a DP/NP complement.

(9) [$_{CP}$ [that he might be wrong]$_i$ [$_{CP}$ pro $_i$ [$_{TP}$ he didn't think of t $_i$]]]

The pro-approach is recently revived and elaborated by Alrenga (2005). This section examines whether the Koster-Alrenga approach can handle the other types of movement paradoxes, concluding that it is empirically inadequate in view of the type 4 movement paradox and agreement phenomenon.

2.1 The Non-DP/NP Subject as a Topic

The type 2 movement paradox is characterized by the occurrence of so-called non-DP/NP subjects. If the non-DP/NP subjects are not subjects but topics, type 2 can be treated just like type 1. If the non-DP/NP at the edge of a sentence is a topic and if the real subject is a pro, then it does not come as a surprise that there is a movement paradox with regard to NP-movement;

(10) [$_{CP}$ topic$_i$ [$_{CP}$ pro $_i$ [$_{TP}$ t $_i$ …verb t $_i$]]]
        |_____||_____|

This is exactly what Koster (1978) and Alrenga (2005) claim. They present several phenomena in support of their claim that the sentential subjects are Topic Phrases. First of all, Koster observes that sentential subjects and Topic Phrases show the same pattern with regard to subject-aux inversion in that they do not allow subject-aux inversion.

(11) *Did the book$_i$ Mary give t$_i$ to John?
(12) a. *Did [that John showed up] please you?
     b. Did it please you that John showed up? (Koster 1978)

Second, both of them do not permit A'-movement. Sentences (13a-b) below illustrate that A'-movement across either a Topic Phrase or a sentential subject is not permitted.

(13) a. *John$_j$, [that the Giants lost the world series] shouldn't have bothered t$_j$.
     b. *John$_j$, the book$_i$ I gave t$_i$ to t$_j$.   (Alrenga 2005: 179)

Third, both of them, unlike DP/NP-subjects, are not permitted inside islands. They cannot occur inside adjunct clauses, clausal complements of nouns, clausal topics or other clausal subjects, whereas DP/NP subjects can occur in these positions, as exemplified by (14-16).

(14) a. ?*Mary is unhappy because for her to travel to Tahiti is no longer
        necessary.

b. ?*Although that the house is empty depresses you, it pleases me.

c. ?*Jim raised the possibility that for the house to be destroyed would upset you.

d. ?*That for us to smoke would bother her, I didn't expect.

e. ?*That for us to smoke bothers her is obvious. (Alrenga 2005: 178)

(15) a. *Mary is unhappy because her trip to Tahiti, I've had to cancel.

b. ?*Although Mary, this may depress, it pleases me.

c. *John raised the possibility that Mary, your antics would upset.

d. *That Mary, our antics would upset, I didn't expect.

e. *That Mary, your antics will upset is obvious. (Alrenga 2005: 179)

(16) a. Mary is unhappy because her trip to Tahiti is no longer necessary.

b. Although the house's emptiness depresses you, it pleases me.

c. Jim raised the possibility that the house's destruction would upset you.

d. That our smoking would bother her, I didn't expect.

e. That our smoking bothers her is obvious. (Alrenga 2005: 178)

Fourth, they can only occur in the complements of bridge verbs. Sentences (17-19) below show that they are permitted in the complements of bridge verbs, but not in the complements of non-bridge verbs.

(17) a. I {think/said/believe} that for us to smoke really bothers her.

b. ?*I regret that for us to smoke bothers her so much.

c. ?*Mary wishes that for us to smoke bothered her more than it did.
(Alrenga 2005: 178)

(18) a. I {think/said/believe} that our smoking really bothers her.

b. I regret that our smoking bothers her so much.

c. Mary wishes that our smoking bothered her more than it did.
(Alrenga 2005: 178)

(19) a. Mary {thinks/said/believes} that John, the article really bothered.

b. ?*I regret that Mary, my antics upset as much as they did.

c. ?*Mary wishes that John, the article bothered more than it did.
(Alrenga 2005: 179)

Koster and Alrenga argue that all the above data suggest that sentential subjects are Topic Phrases, so that they are base-generated in the topic position.[1]

---

[1] They assume that Topic Phrases are base-generated at the edge of a sentence.

2.2 Alrenga's Generalization

According to the Koster-Alrenga approach, it is not Topic Phrase but pro that is selected by the main predicate. If this approach is correct, it is predicted that since the grammatical category of pro is a DP/NP, sentential subject constructions are permitted only if the main predicate can select a DP/NP. Alrenga (2005) argues that it is indeed the case.

(20) Alrenga's generalization:

A passive verb may appear with a sentential subject only if the position of the gap is the one in which a DP is licensed by the verb's active form.

    Alrega notes that the verbs like *hoped/felt/wished/insisted/reasoned* do not license a DP/NP complement, and interestingly, the passive counterpart of those predicates do not permit CP raising and instead, *it* is inserted.

(21) Most baseball fans hoped/felt/wished/insisted/reasoned that the Giants
      would win the World Series.

(22) *Most baseball fans hoped/felt/wished/insisted/reasoned that.

(23) *That   the   Giants   would   win   the   World   Series   was
      hoped/felt/wished/insisted/reasoned (by most baseball fans).

(24) It was hoped/felt/wished/insisted/reasoned that the Giants would win the
      World Series. (Alrenga 2005: 183)

The predicates *hope for* and *wish for*, in contrast to *hope* and *wish*, take a DP/NP-complement only: they cannot take a clausal-complement. Interestingly, a well-formed sentence is generated when the subject is clausal.

(25) *John {hoped for, wished for} that the Giants would win the world series

(26) That the Giants would win the world series seems to have been {hoped for,
      wished for} (by most baseball fans)} (Alrenga 2005: 198)

The data in (21-26) are all compatible with the generalization that the clausal subject is permitted only if the predicate selects a DP/NP complement.

    Let us now examine how the pro-approach can deal with the movement paradoxes. We have already seen that the type 1 paradox can be explained. Now, type 2 paradox can be explained in the same fashion. Sentence (3a), rewritten here as (27), is excluded because *capture* cannot c-select a CP.

(27)   *This theory captures that languages are learnable.

Under the pro-approach sentence (3b), repeated here as (28a), would be represented as (28b).

(28) a. That languages are learnable is captured by this theory.

  b. [$_{CP}$ [that languages are learnable]$_i$ [$_{CP}$ pro $_i$ [$_{TP}$ t $_i$ is captured t$_i$ by this theory]]]

In (28b) the c-selection of *capture* is satisfied, since pro is an NP. Therefore, (28a) is grammatical. Type 3 paradox can also be explained; sentence (5b), here numbered as (29a), would be represented as (29b) but the representation is ill-formed because *think* does not select a DP/NP-complement.

(29) a. *[That he might be wrong], he didn't think.

  b. [$_{CP}$ [that he might be wrong]$_i$ [$_{CP}$ pro$_i$ [he didn't think t$_i$]]]

The Koster-Alrenga approach can deal with the types 1, 2 and 3 movement paradoxes, but it fails to explain type 4. As illustrated in (30), the verb *hold* licenses a DP/NP-complement, but the passive verb *held* cannot occur with a clausal subject. There seems to be nothing wrong with the representation in (31b), but (7b), numbered here as (31a), is ungrammatical.

(30) Theodore never held that/such a view.

(31) a. *That the chimp was immortal was held by Theodore.

  b. [$_{CP}$ [that the chimp was immortal]$_i$ [$_{CP}$ pro $_i$ [$_{TP}$ t $_i$ was held t $_i$ by Theodore]]]

In short, the type 4 movement paradox runs counter to Alrenga's generalization, which suggests that the Koster-Alrenga's approach is on the wrong track. In addition, there is one more serious problem with the Koster-Alrenga's approach. The pro-approach fails to capture the agreement phenomenon of the sentential subject constructions.

(32) a. That he will resign and that he will stay in office are/*is at this point equally likely.

  b. It is/*are at this point equally likely that he will resign and he will stay in office.

Alrenga (2005) argues that CP has no number feature on the ground that (32b) is ungrammatical when the copula is in the form of *are*. The post-copular NPs agree with the copular, as illustrated by (33), but in (32b) the copula does not agree with the conjoined CP.

(33) There are/*is three dogs in the yard.

This consideration leads Alrenga to claim that CP is numberless. On the other hand, with a view to explaining the phenomenon in (32a), he claims that "the [+plural] feature of the null DP is determined by its anaphoric link with the

coordinate CP, which is base-generated in the topic position". This claim runs into contradiction with his previous claim that CP has no number feature. He argues that CP is numberless on the basis of (32b) and simultaneously he argues that it has a number feature on account of (32a). It seems that there is no way to explain the agreement phenomenon under the pro-approach.

2.3 The Trace Approach

Let us now consider another approach which is similar to the pro-approach in that it avails itself of an empty category. We have seen that the Koster-Alrenga approach makes use of pro. Stowell (1981) and Webelhuth (1992) utilize another type of empty category—trace. Stowell (1981) proposes that the *that*-clause is not base-generated in the topic position but it is generated in the thematic position and moves to the subject position. He goes on to argue that it cannot stay in the subject position and undergoes further movement to the topic position, since it cannot be assigned Case.

(34) [[that pigs can fly]$_i$ [$_{TP}$ t$_i$ seems to t$_i$ be surprising]]

In this analysis the *that*-clause and its trace are asymmetric in the sense that the trace can be assigned Case, even though the clausal subject cannot. This is reminiscent of the pro-approach. Under the pro-approach the movement paradoxes arise from the asymmetry between the non-DP/NP topic and the pro associated with the topic. In the trace approach the asymmetry between the clause and its trace is responsible for the movement paradoxes. Stowell assumes that traces are NPs, since they can be assigned Case. Webelhuth (1992) put forth a similar idea. He proposes that traces have no feature but they are given the default grammatical category [+noun]. The trace-as-an-NP approach straightforwardly captures Alrenga's generalization. *Think of* and *capture* take a DP/NP-complement, but not a CP-complement. The sentences in (35a) and (36a) are not grammatical, for they can co-occur with a *that*-clause only. On the other hand, (35b) and (36b) are grammatical because the trace is an NP.

(35 = 1)   a. *He didn't think of [that he might be wrong].
           b. [That he might be wrong], he didn't think of [$_{NP}$ t$_i$].

(36 = 2)   a. *This theory captures [that languages are learnable].
           b. [That languages are learnable]$_i$ is captured [$_{NP}$ t$_i$] by this theory.

By contrast, *think* c-selects a CP, not a DP/NP. This c-selection is satisfied in (37a), but not in (37b). Therefore, there is a contrast in grammaticality between

them.[2]

(37 = 5)   a. He didn't think [that he might be wrong].
        b.*[That he might be wrong], he didn't think [$_{NP}$ t$_i$].

Just like the pro-approach, the trace-approach manages to explain types 1, 2 and 3, but it cannot handle type 4. The predicate *hold* can take a DP/NP-complement, and so (39b) satisfies the c-selection of *hold*. Surprisingly, however, it is ill-formed.[3]

(38 = 30) Theodore never held that/such a view.

(39 = 7)   a. Theodore held [that the chimp was immortal].
        b. *[That the chimp was immortal] was held [$_{NP}$ t$_i$] by Theodore.

Furthermore, there are many other problems with this approach. First, the notion 'trace' is not compatible with the No Tampering Condition. In accordance with the No Tampering Condition traces must not be created in the course of a derivation. However, the trace-approach is heavily dependent on the notion 'trace', which makes it conceptually unattractive. Second, it is unclear how to explain the agreement phenomena under this approach. We have to say that the trace in (40a) is plural, while the trace in (40b) is singular.

(40 = 24) a. [That he will resign and that he will stay in office]$_i$ t$_i$ are/*is at this
        point equally likely.
        b. [That he will resign]$_i$ t$_i$ *are/is at this point likely.

But there is no principled ground that the trace of the conjoined clause must be plural, while that of the non-conjoined clause must be singular; the *that*-clause, to which the trace is linked, has no Φ-features and if so, neither can the trace have Φ-features. To recapitulate, type 4 movement paradox and the agreement phenomenon are left unexplained if we make use of an empty category like a pro or a trace.

3.   A New Account

Dubinksy and Davies (2006) argue that clausal subjects are noun phrases on the ground that the plural form *are* is required in (40a), whereas the singular form is

---

[2] The movement approach assumes that the c-selectional restriction is checked at PF.
[3] The base-generation approach advocated by Koster and Alrenga and the movement approach proposed by Stowell and Webelhuth are quite similar in that a null constituent takes care of the c-selection of a predicate and Case. The only difference between them is whether *that*-clauses are base-generated in the topic position or they are generated in a thematic position and moves to the topic position.

required in (40b). Furthermore, they point out that clausal subjects and topic phrases show a different pattern with regard to subject-aux inversion. According to them, many native speakers of English judge that subject-aux inversion is permitted even when the subject is a non-DP/NP.

(41) a. Would <u>to leave early</u> really reflect poorly on us all?

b. When does <u>to let the sleeping dogs lie</u> strike you the best course of action?

(42) a. To whom is <u>that pigs can fly</u> most surprising?

b. Is <u>that I am done with the homework</u> really amazing?
(Dubinsky and Davies 2006:7)

The phenomena in (40a-b) and (41-42) can be explained if clausal subjects are NPs. If *that*-clauses can be NPs, however, there is no way to account for (43a).

(43) a. It *are/is at this point equally likely [that he will resign and that he will stay in office].

b. There are/*is three dogs in the yard.

The copula agrees with the post-copular DP/NP, as illustrated in (33), numbered here as (43b). If the *that*-clause in (43a) is an NP, it is incorrectly predicted that *are* can occur in the sentence. Furthermore, under the assumption that the *that*-clause can be used as an NP, it is not possible to explain the ungrammaticality of the sentences like (44).

(44)   *John didn't think of that he might be wrong.

What we want to say is that the *that*-clause at the left edge of a sentence is an NP, but the one in the thematic position is not. This sounds quite strange. However, this is exactly what this article claims. This section shows that this seemingly contradictory claim follows from the fact that the constituent headed by Topic is a DP/NP regardless of its original grammatical category and it must undergo topicalization.

3.1 The Properties of Topic Phrases and Movement Paradoxes

I assume that non-DPs/NPs have no Φ-features. I claim, however, that even non-DPs/NPs come to have Φ-features if they occur with a topic head—if they are Topic Phrases. The Topic Phrase has the structure in (45), where it consists of the null topic head and the content of the topic.

(45)   The structure of Topic Phrase: $[_{\text{Topic P}} \emptyset_{[+\text{Topic}, +\text{nominal}]} [_{\text{XP}} \ldots]]$

Kwang-sup Kim

Under this analysis the *that*-clause in (46) is analyzed as (47).

(46)   That he will go to New York soon is obvious.

(47)   [$_{TopicP}$ Ø $_{[+Topic, +nominal]}$ [$_{CP}$ that he will go to New York soon]]

Since the null topic head is an NP, it has Φ-features. Accordingly, the Topic Phrase can move to the SPEC of TP even if its original grammatical category is not DP/NP. After moving into the subject position, the Topic Phrase, which has a strong topic feature, must undergo a further movement—topicalization.

(48) [$_{TopicP}$ [$_{TopP}$ Ø [$_{CP}$ that he will go to NY soon]]$_i$ Topic [$_{TP}$ t$_i$ is [t$_i$ obvious]]]
|_____||___|

In short, the topic head has two important features: [+noun] and [+topic], and the former enables Topic Phrases to have Φ-features, which can trigger A-movement, and the latter forces them to undergo topicalization, since it is a strong feature. Topics must be referential, and DP/NP, not CP, is the category that can be referential. So it is not surprising that the topic phrase is DP/NP, not CP.[4]

Now, the type 1 movement paradox can be explained straightforwardly. The type 1 predicates like *think of* subcategorize for an NP, and Top Phrases are NPs.

(49) Type 1 Predicates: [____, DP/NP]

Accordingly, the predicate can take a Topic Phrase as its complement. The Topic Phrase cannot stay in the thematic position but it must move into the SPEC of Topic.

(50) a. [Topic [$_{TP}$ He didn't think of [$_{Topic\ P}$ Ø [$_{CP}$ that he might be wrong]]]]:
                                                        Topicalization
     b. [[$_{Topic\ P}$ Ø [$_{CP}$ that he might be wrong]]$_i$ Topic [$_{TP}$ he didn't think of t$_i$]]
                     |_____|

Therefore, a well-formed sentence is generated only if the *that*-clause is topicalized. Put differently, the type 1 movement paradox arises because the Topic Phrase must undergo topicalization.

The type 2 paradox can be explained in a similar fashion. The type 2 predicates subcategorize for a DP/NP only, so that it is impossible to generate (3a) and (52), in which CP is a complement.

---

[4] Section 5 shows that there is cross-linguistic evidence that CP cannot be a topic unless it is nominalized.

(51)   Type 2 predicates: [___, DP/NP]

(52)   *This theory captures [$_{CP}$ that languages are learnable].

On the other hand, the type 2 predicates can take as their complement a Topic phrase, since it is an NP. The Topic Phrase, which has Φ-features, undergoes A-movement, and then it undergoes further A'-movement—topicalization.

(53) a. [[$_{Topic P}$ Ø$_{[+nominal]}$] [$_{CP}$ that languages are learnable]]$_i$ [$_{TP}$ t$_i$ is captured t$_i$ by this theory]

   b. That languages are learnable is captured by this theory.

Therefore, (3b), rewritten as (53b), is grammatical, even though (52) is not. Let us now consider type 3. The predicates like *think* subcategorize for a clause only.

(54)   Type 3 predicates: [___, clause]

Accordingly, Topic Phrase cannot be a complement of the type 3 predicates, since it is not a clause but an NP. For instance, (55a) is in violation of the c-selectional requirement of *think*.

(55) a. *[$_{TP}$ He didn't [think [$_{Top P}$ Ø$_{[+nominal]}$ [$_{CP}$ that he might be wrong]]]]]

   b. [$_{TP}$ He didn't [think [$_{CP}$ that he might be wrong]]]

The representation in (55b) is well-formed, but the CP *that he might be wrong* cannot undergo topicalization, for it has no topic feature. Therefore, there is no way to topicalize the *that*-clause when the predicate is *think*.

   Let us finally turn to type 4. I propose that the type 4 predicates only s-select a proposition.

(56)   Type 4 predicates: [___, proposition]

The proposition can be realized in the form of DP/CP as well as CP. Sentence (57) is a case in point. The DP *such a view* in (57) is interpreted as a proposition: (57) is interpreted as 'John never held **such a view to be true**'.

(57) John never held such a view.

   On the other hand, the Topic Phrase cannot satisfy the s-selectional restriction of the type 4 predicate, for it cannot denote a proposition. CP denotes a proposition but the null topic head turns the proposition to a referential entity. A topic is what the sentence is about, and we can talk about a referential entity only. This means that the topic must denote a referential entity. Since the predicate *hold* does not s-select a referential entity, the representation in (58a) must be excluded. It seems that (58b) is the only available option, but it cannot be a source for (59), for the *that*-clause has no Φ-features and so cannot undergo A-movement. Therefore, there is no way to generate (59).

(58) a. *[TP was [held [Topic P that the chimp was immortal] by John]]

     b. [TP was [held [CP that the chimp was immortal] by John]]

(59)   *That the chimp was immortal was held by John.

Postal (2004) observes that besides *hold*, there are many other predicates which can take NP-complements, but do not permit clausal subjects.

(60) a. She felt/found/heard that the chimp was intelligent.
     b. *That the chimp was intelligent was felt/found/heard by Stan.
     c. She felt/found/heard that.

(61) a. Stella couldn't conceive that the chimp was intelligent.
     b. *That the chimp was intelligent couldn't be conceived by Stella.
     c. Stella couldn't conceive that/such a thing/anything of the sort.

(62) a. Irma intended that you see the results.
     b. *That you see the results was intended by Irma.
     c. Irma didn't intend that/any such thing.

(63) a. This means that we will be indicted.
     b. *That we will be indicted is meant by this.
     c. This has got to mean that/something.

(64) a. Aristotle pledged that he would resign.
     b. *That he would resign was pledged by Aristotle.
     c. Aristotle did not pledge that/anything.   (Postal: 2004)

Interestingly, all the nominal complements in the c-sentences of (60-64) refer to the content of a proposition—the content of what she felt/found/heard/, the content of what Stella couldn't conceive, and so on. This is evidenced by the fact that the question starting with the nominal *wh*-word *what* is answered by a proposition in those constructions.

(65) A: What did she feel/find/hear?
     B: She felt/found/heard that that the chimp was intelligent.

(66) A: What did she conceive?   B: She conceived that the chimp was
                                intelligent.

(67) A: What did Irma intend?   B: He intended that you see the results.

(68) A: What does this mean?   B: This means that we will be indicted.

(69) A: What did Aristotle pledge?   B: He pledged that he would resign.

In all the examples above the content of *what* is specified as a proposition. This suggests that all those predicates s-select a proposition, whether they take CP-complements or NP-complements. This is further supported by the fact that the nominal complements of the type 4 predicates cannot be topicalized in a positive context. All the sentences below are awkward, and barely acceptable.[5]

(70) a. ??That, Theodore held.   b. ??That, she felt/found/heard.
      c. ??That/Such a thing, Stella could conceive.   d. ??That, Irma intended.
      e. ??That/Something, this has got to mean.   f. ??That, Aristotle pledged.

This follows if even the nominal complements denote a proposition and the proposition cannot be a genuine topic. In short, the type 4 paradox originates from the fact that the type 4 predicates s-select a proposition, but the proposition-denoting XP cannot undergo topicalization. To sum up, all the four types of movement paradoxes can be explained by the claim that Topic Phrase is an NP and must undergo topicalization, along with the argument structure of each predicate.[6] Another important point is that movement paradoxes undermine neither the movement approach, nor the copy theory of movement.

3.2 Differences between CP and PP

The gist of the claim made here is that the topicalized non-DPs/NPs are nominals. For instance, CPs and PPs turn into DPs/NPs when they are headed by the null topic. Interestingly, PPs differ from CPs in that even non-topicalized PPs behave like DPs/NPs in certain contexts. As sentences (71-72) below show, PPs can occupy the complement position of transitive predicates and they can undergo passivization just like DPs/NPs.

---

[5] Fronting of the nominal complements seems to be permitted in the negative context.

(i)   a. That, Theodore never held.   b. That, she didn't feel/find/hear.
      c. That/Such a thing, Stella could not conceive.   d. That, Irma did not intend.
      e. That, Aristotle did not pledge.

The fronted nominal complements are not neutral topics but contrastive topics or focuses; they need to be inside the scope of negation but the topics must be under the scope of negation. See Section 5 for a discussion of the relation between the topic phrase and the positive/negative context.

[6] It is noteworthy that both the pro-approach and the trace-approach cannot explain the type 4 movement paradox; CP is propositional and so the empty category that is anaphorically linked with the CP must be propositional. If the empty categories can be propositional, it is incorrectly predicted that the representation in (i) is well-formed.

(i) That the chimp was immortal was held {pro, t} by John.

Kwang-sup Kim

(71) a. They discussed after the holidays.
　　　b. After the holidays was discussed by everyone in the office for the entire month of February.　　(Dubinsky and Davies (2006: 21-22))

(72) a. Where will the new discotheque be built?
　　　b. I don't know, but the council rejected behind the church. (Aarts 2001: 76)

This does not come as a surprise in view of the fact that PPs can be a complement of another preposition.

(73) [PP from [PP inside the church]]

The data in (71-73) suggest that there are two types of PPs: referential PPs and relational PPs. In (74a) the preposition *on* is a two-place predicate, taking *the table* as its first argument and *apples* as its second. The preposition *in* in (75b) is also a two-place predicate in that its first argument is *the park*, and its second argument is an event.

(74) a. Apples are on the table.　　b. John met Mary in the park.

(75) a. on (the table, apples)　　　b. in (e, the park)

On the other hand, *under the stairs* is not relational in (76a) and instead, it refers to a certain place. In that sense it is referential.

(76) a. Under the stairs was a safe area to be during the war.
　　　b. Outside the fridge is not a good place to keep milk.
　　　c. After Saturdays would be a good time to go away for a few days.
　　　d. Between eleven and midnight suits me alright. (Aarts (2001: 73))

Let us look into the difference in argument structure between the referential PPs and the relational PPs. It is well-known that the NPs like *a genius* can be used as an argument or as a predicate.

(77) a. John is a genius.　　　　b. John met a genius.

　　The difference between the predicative NP and the argument NP lies in whether it can satisfy the thematic role of the head noun. The predicative NP cannot satisfy the thematic role of *genius* and it is transmitted to *John*, whereas the argument NP can satisfy the thematic role of *genius*; so-called vertical binding takes place when the NP is an argument. The thematic role Theme is assigned to *John* in (78a), but the thematic role Reference is satisfied by the NP *a genius* in (78b).

(78) a. [be [John$_i$ a genius$_{(\text{theme } i)}$]]　　b. [meet [$_{NPi}$ a genius$_{(\text{Reference } i)}$]]
　　　　　　|_____|　　　　　　　　　　　　　　　　|_____|

Likewise, I propose that the relational PP and the referential PP are different in terms of argument structure. The thematic role of the relational PP must be transmitted to its sister, while the thematic role of the referential PP is satisfied by its maximal projection—PP.

(79) a. [the cat$_x$ [under$_{\text{(Theme x, Location y)}}$ the table$_y$]]
    b. [$_{PPx}$ under$_{\text{(Reference x, Location y)}}$ the table$_y$]]

Referential PPs, just like NPs, can have Φ-features, and so it is not surprising that they agree with tense and permit subject-aux inversion when they occur in the subject position

(80) Under the table and inside the closet are not good places to hide them.

(81) Is under the table a good place to hide them?

To recapitulate, there are two types of nominal PPs: originally referential PPs and topicalized relational PPs. They are slightly different in their syntactic behavior: the topicalized PPs, unlike the originally referential PPs, cannot be involved in agreement even though they are NPs. The PPs in (82a-b) cannot agree with the copula, since they are not originally referential PPs but topicalized predicative or relational PPs.

(82) a. Under the table are lots of apples.   b. In the yard are many dogs.

Let us attempt to generate sentence (82a). The relational PP can assign a thematic role even if it is headed by the nominalizer—the topic head, for thematic roles can be percolated via a nominal head. For instance, the thematic role of the verb *claim* is percolated even if it is nominalized.

(83) a. John claimed$_{\text{(agent, theme)}}$ that he is a genius.
    b. John's [$_N$ ø [$_V$ claim]]$_{\text{(agent, theme)}}$ that he is a genius.

(84) a. [$_N$ ø [$_V$ claim$_{\text{(agent, theme)}}$]]: percolation of thematic roles
    b. [$_N$ ø [$_V$ claim$_{\text{(agent, theme)}}$]] $_{\text{(agent, theme)}}$

Likewise, the relational PP *under the table* can maintain the thematic structure even after it is selected by the null topic head. Accordingly, *[$_{TopP}$ ø [under the table]]* can take *lots of apples* as its argument, forming a small clause, as shown by (79a).

(85) a. [lots of apples [$_{TopP}$ ø [under the table]]]: merger of *be* and T
    b. [T [be [lots of apples [$_{TopP}$ [under the table]]]]]:
        Agreement with *lots of apples* and raising of *[$_{TopP}$ [under the table]]*
    c. [$_{TP}$ [$_{TopP}$ [under the table]] T [be [lots of apples [$_{TopP}$ [under the table]]]]]:
                                    Topicalization

d. [$_{CP}$ [$_{TopP}$ [under the table]] [$_{TP}$ ~~[$_{TopP}$ [under the table]]~~ T [be [lots of apples ~~[$_{TopP}$ [under the table]]~~]]]]

There is a disparity between agreement and movement. The derivation in (85a-d) illustrates that T agrees with the argument *lots of apples*, not with the predicate *[$_{TopP}$ ø [under the table]]*, but the SPEC of T can be filled by the Topic Phrase. It is because only an argument, not a predicate, can determine the inflectional ending of T, but the SPEC of T can be occupied even by a non-argument as far as locality conditions are met. For instance, in *there*-constructions the post-copular NP agrees with T, but it does not move into the SPEC-T and rather *there*-insertion takes place.

(86) a. [$T_{[\alpha \text{ person}, \beta \text{ pl}]}$ be dogs $_{[3rd \text{ person}][+pl]}$ in the yard]: agreement with *dogs*
    b. [$T_{[3rd \text{ person}][+pl]}$ be dogs$_{[3rd \text{ person}][+pl]}$ in the yard]: merger of *there*
    c. [there T be dogs in the yard]

In (85a) the predicative Topic Phrase and its argument c-command each other, and the small clause is not a phase. So it is possible for the predicative Topic Phrase to move into the SPEC-T, even though it cannot agree with T. To recapitulate, there are two types of PPs: the relational PPs and the referential PPs. The referential PPs pattern like NPs, and even the relational PPs can turn into NPs if they are headed by Topic, but the topicalized PPs and the originally referential PPs differ in that the former type, unlike the latter, cannot be involved in agreement.[7]

6. Conclusion

There are many instances in which movement turns an ungrammatical sentence into a grammatical one and a grammatical sentence into an ungrammatical one. These movement paradoxes arise from the fact that there is an asymmetry between the dislocated constituent and the constituent in situ. The empty category approaches like the pro-approach and the trace approach assume that the asymmetry originates from the fact that the two constituents are different entities: the constituent at the edge of a sentence is a CP, while the constituent in the complement position is a nominal empty category. This article has argued that those empty category approaches fail to explain the type 4 movement paradox and agreement phenomenon: they cannot be explained if we assume that the two constituents are different. I have put forth the idea that the dislocated constituent

---

[7] In (2a-b) *on the wall* is a relational PP so that it can behave like an NP only if it is attached by the topic head.

and the constituent in situ are identical and have the same grammatical category—DP/NP, which in turn follows from the natural claim that Topic Phrase undergoes topicalization and Topic Phrases are DPs/NPs. Under this analysis the two constituents are the same, since they are copies of the same constituent. The only difference between the two copies lies in whether the strong topic feature is checked or not. The constituent with a strong topic feature must not be pronounced, and as a consequence, only the copy at the topic position can be pronounced, which gives rise to all the four types of movement paradoxes.

(87) [$_{TopicP}$ [$_{Topic\ P}$ topic$_{[+strong\ topic]}$ [$_{CP}$ …]]]$_i$ Topic….. [$_{Topic\ P}$ topic$_{[+strong\ topic]}$ [$_{CP}$ …]]]

To summarize this article, movement paradoxes are not paradoxes but they follow from the fact that Topic Phrase is a DP/NP and it must undergo topicalization, since it has a strong topic feature.

References

Aarts, Bas. 2001. *English syntax and argumentation*. Hampshire, UK: Palgrave.

Alrenga, Peter. 2005. A sentential subject asymmetry in English and its implications for complement selection, *Syntax* 8-3:175-207.

Arthur, J.-M. 1992. Iterated CPs and embedded topicalization. *Linguistic Inquiry* 23:329-336.

Belletti, A. & L. Rizzi. 1988. Psych-verbs and θ-theory. *Natural Language & Linguistic Theory* 6:291-352.

Bresnan, Joan. 1994a. Category Mismatches. Presented to the plenary session of the 25th Annual Conference on African linguistics at Rutgers University on March 26, 1994.

Bresnan, J. 1994b. Locative inversion and the architecture of universal grammar. *Language* 70:72-131.

Bresnan, J. 1995. Category mismatches. In *Theoretical Approaches to African Linguistics*, ed. by A. Akinlabi, 19-46. Trenton, NJ: Africa World Press.

Chomsky, N. 1995. *The Minimalist Program*. Cambridge: MIT Press.

Chomsky, N. 2001a. Derivation by Phase. In *Ken Hale: A life in language,* ed. by Michael Kenstowicz, *1-52*. Cambrdige: MIT Press.

Chomsky, N. 2001b. Beyond Explanatory Adequacy. *MIT Occasional Papers in Linguistics 20*. Cambridge, MA: MITWPL.

Cinque, G. 1990. Ergative adjectives and the lexicalist hypothesis. *Natural Language and Linguistic Theory* 8:1-39.

Culicover, Peter W and Michael S.Rochemont. Extraposition and the complement principle. *Linguistic Inq*uiry 21: 23-47.

Davies, W. D. and S. Dubinsky. 2001. Functional architecture and the Distribution of subject predicates. In *Objects and other subjects: Grammatical Functions, functional categories and configurationality*, ed. by William Davies and Stanley Dubinsky, 247-279. Dordrecht: Klwuer.

Kwang-sup Kim

Davies, W. D. and S. Dubinsky. 2006. Sentential (and non-nominal) subjects in English. *Studies in Modern Grammar* 45: 1-32.

Delahunty, G. 1983. But sentential subjects do exist. *Linguistic Analysis* 12: 379-398.

den, Dikken, Marcel. 2006. *Relators and linkers: the syntax of predication, predicate inversion and copulas*. Cambridge, MA.: MIT press.

Emonds, J. 1976. *A transformational approach to English syntax*. New York: Academic Press.

Grimshaw, J. 1979. Complement selection and the lexicon. *Linguistic Inquiry* 10:279-326

Higgins, F. R. 1973. On J. Emonds's analysis of extraposition. In *Syntax and semantics* 2, ed. by J. Kimball, 149-195. New York: Seminar Press.

Hooper, J. B. & S. A. Thompson. 1973. On the applicability of root transformations. *Linguistic Inquiry* 4:465-497.

Iatridou, S. and A. Kroch. 1992. The licensing of CP-recursion and its relevance to the Germanic verb-second phenomenon. *Working Papers in Scandinavian Syntax* 50: 1-24.

Iwakura, K. 1976. Another constraint on sentential subjects. *Linguistic Inquiry* 7:646-652.

Jacobson, P. 1992. The lexical entailment theory of control and the *tough*-construction. In *Lexical matters*, ed. by I. Sag & A. Szabolcsi, 269-299. Stanford, Calif.: CSLI Publications.

Jaworska, Ewa. 1986. Prepositional phrases as subjects and objects. *Journal of Linguistics* 22:355-374.

Johnson, Kyle. 1991. Object positions. *Natural Language & Linguistic Theory* 9:577-636.

Kim, Kwang-Sup. Subject-Predicate Agreement, Dummy Subject and EPP. *Proceedings of SICOL* 2006: 194-206.

Koster, J. 1978. Why subject sentences don't exist. In *Recent transformational studies in European languages*, ed. by S. J. Keyser, 53-64. Cambridge, Mass.: MIT Press.

Kuno, Susumu. 1973a. Constraints on internal clauses and sentential subjects. *Linguistic Inquiry* 4:363-385.

Newmyer, Frederic. Non-DP subjects and objects in English and their theoretical implications. *Proceedings of SICOL* 2006: 207-214.

Pesetsky, David. 1982. Paths and categories. Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge.

Postal, Paul. 2004. *Skeptical Linguistic Essays*. Oxford: Oxford University Press.

Rizzi, Luigi. 1990. *Relativized minimality*. Cambridge, MA.: MIT press.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of grammar*, ed. by L. Haegeman, 281-337. Dordrecht: Kluwer.

Ross, J. R. 1967. Constraints on variables in syntax. Doctoral dissertation, MIT.

Stowell, Tim. 1981. Origins of phrase structure. Doctoral dissertation, MIT.

Tsai, W-T. 1995. Visibility, complement selection and the case requirement of CP. *Journal of East Asian Linguistics* 4:281-312.

Webelhuth, G. 1992. *Principles and parameters of syntactic saturation*. Oxford: Oxford U. P.

Williams, Edwin. 1981. Argument structure and morphology. *Linguistic Review* 1:81-114.

Williams, Edwin. 1994. *Thematic Structure in syntax*. Cambridge, Mass.: MIT Press.

On the Scope Properties of Nominative Phrases in Japanese[*]

Mamoru Saito

Nanzan University and University of Connecticut

Abstract

The scope properties of nominative objects in Japanese have been discussed extensively since Tada (1992) and Koizumi (1995, 1998). The basic observation is that they take scope at a higher position than accusative objects. The papers just cited argue that this is because nominative objects are raised by overt A-movement to a relatively high position in the structure for the purpose of Case checking. This paper builds on these works and argues that the relevant movement is covert A'-movement. One of the consequences of this analysis is that the EPP-feature of T attracts the closest NP independently of nominative Case licensing, a hypothesis proposed and discussed in Ura (1999). It is shown that this leads to a solution for the long-standing puzzle regarding the strict A-property of short scrambling to the post-subject position.

1. Introduction

The main purpose of this paper is to discuss the licensing mechanism of the nominative particle *-ga* in Japanese. I will argue in particular that nominative objects move covertly to the outer Spec of TP for the purpose of Case licensing. Then, I will examine the consequences of this proposal for the nature of the EPP requirement of T as well as for the strict A-property of short scrambling to the post-subject position.

The particle *-ga* is referred to as the nominative Case marker. But it is well known that its distribution is wider than the English nominative. For example, Japanese allows multiple subjects and in this case, all the subjects are accompanied by *-ga*. This is shown by (1) from Kuno (1973):

(1) [TP Bunmeikoku -ga [TP dansee-ga [TP heikinzyumyoo -ga   mizika-i]]]
        civilized-country-NOM male-NOM average-life-span-NOM short
      'It is civilized countries where the male population has a short life-span'

In addition, objects of stative predicates are marked by *-ga*, as pointed out in classical works like Kuroda (1965) and Kuno (1973).

Mamoru Saito

(2) a. Taroo-ga   kantongo -o   hanas-u (koto)
               -NOM Cantonese-ACC speak   (fact)
         '(the fact that) Taroo speaks Cantonese'

    b. Taroo-ga   /-ni   kantongo -ga   wakar-u   (koto)
               -NOM/-DAT Cantonese-NOM understand fact
         '(the fact that) Taroo understands Cantonese'

    c. Taroo-ga   /-ni   kantongo -o   /-ga   hanas-e-ru (koto)
               -NOM/-DAT Cantonese-ACC/-NOM speak-can (fact)
         '(the fact that) Taroo can speak Cantonese'

As (2a) shows, the accusative *-o* accompanies the object when the predicate is non-stative. However, *-ga* is employed for stative predicates as in (2b). The interesting case that has often been discussed in the literature is (2c). The predicate consists of two verbal morphemes, the non-stative *hanas* 'speak' and the stative *-(rar)e* 'can'. In this case, the object can be marked by either *-ga* or *-o*.[1]

It has been assumed that the stative verbal suffix *-(rar)e* 'can' plays a role in the nominative marking of the object in examples like (2c). One of the concrete hypotheses for this was proposed in Tada (1992) on the basis of the scope facts in (3), originally observed by Sano (1985).

(3) a. Kiyomi-wa   migime -dake-o   tumur-e-ru (can > only)
               -TOP right-eye-only -ACC close -can
         'Kiyomi can wink with her right eye'

    b. Kiyomi-wa   migime -dake-ga   tumur-e-ru (only > can)
               -TOP right-eye-only -NOM close -can
         'It is only her right eye that Kiyomi can close'

As shown in (3a), accusative objects take narrow scope with respect to *-(rar)e* 'can'. On the other hand, nominative objects take wide scope as in (3b). Tada (1992) proposed then that nominative objects move to a position within the projection of the verb *-(rar)e* for Case licensing, as illustrated in (4).

(4) [TP Kiyomi [T' [VP [V' [VP right eye only-NOM   close] -can]] -Pres.]]

The landing site is higher than the verbal suffix *-(rar)e* and consequently, they

---

[1] In (2b-c), the subject can be dative only when the object is nominative. The potential verbal suffix assumes the form *-rare* when the final sound of the stem is a vowel and *-e* when it is a consonant.

take wide scope.[2] This movement analysis of nominative objects was further developed in Koizumi (1995, 1998). He pointed out that those objects take wide scope not only over higher predicates, as in (3b), but also over sentential negation. This is shown in (5).

(5) Kiyomi-ga   migime -dake-ga tumur-e   -na-i (koto) (only > not > can)
        -NOM right eye-only -NOM close -can-not   fact
    '(the fact that) it is only her right eye that Kiyomi cannot close'

Accordingly, he proposed that nominative objects undergo overt A-movement to the inner Spec of TP for Case checking, as in (6).

(6) [$_{TP}$ Kiyomi [$_{T'}$ [$_{T'}$ [$_{NegP}$ [$_{VP}$ [$_{VP}$ right eye only-NOM close]-can]-Neg] -Pres.]]]

This analysis achieves a unified treatment of -*ga*: it is always licensed within a projection of T whether it is on the subject or on the object.

    Although the judgment of examples like (5) is fairly clear, a number of linguists, including myself, have not taken these examples as core data because other kinds of examples show different or even contradictory patterns of scope interaction between nominative phrases and negation. More specifically, there are cases where nominative phrases take narrow scope with respect to sentential negation, as observed, for example, in Kuno (1973) and Miyagawa (2003). In this paper, I will first discuss those potentially problematic cases and argue that they are after all consistent with Koizumi's analysis. This, I believe, solidifies the empirical basis of the movement analysis of nominative objects. In Section 3, however, I will raise questions on the specifics of Koizumi's analysis. First, I will argue that the movement in question is not A-movement but is A'-movement. I will discuss some observations in Inoue (1976) and Saito (1982), and show that it is not subject to the locality of A-movement. Secondly, I will argue that the movement is covert as proposed in Ura (1999), building on a discussion in Yatsushiro (1999). If the movement is indeed covert, it must be independent of the EPP requirement of T, as Ura points out. In Section 4, I will

---

[2]   More precisely, Tada (1992) proposes that nominative objects move to the Spec of AgrP projected above the VP headed by -*(rar)e*. As is often noted in the literature, the wide scope construal of accusative objects is marginally possible in examples like (3a). I follow Koizumi (1995) and assume that the object is string-vacuously scrambled to a position higher than   -*(rar)e* in this case. It was also shown by Nomura (2003) that nominative objects can take narrow scope when forced by the context. I assume with him that this is due to reconstruction, which for some reason is a less preferred option in the cases like (3b). See below for some relevant discussion.

Mamoru Saito

consider his analysis of the EPP, which is consistent with this conclusion, and present a supporting argument for it. Section 5 concludes the paper.

2.   The Scope of Nominative Subjects

As noted above, various and sometimes conflicting patterns are observed for the scope of nominative phrases. Those that seem particularly problematic for Koizumi's analysis are the scope of the nominative phrase in existential sentences examined in Kuno (1973) and the scope of subject in scrambling sentences discussed in Miyagawa (2003). I will consider them in turn and argue that the problems they pose are only apparent.

Kuno (1973) examines the correlation of word order and scope relation in existential sentences. It is known since Kuroda (1971) that Japanese exhibits scope rigidity but movement of one quantified NP over another yields scope ambiguity. This is illustrated in (7) and (8).

(7) Dareka -ga   daremo -o   aisitei-ru
      someone-NOM everyone-ACC love
      'Someone loves everyone' (some > every, *every > some)

(8) a. Daremo -o$_i$ dareka -ga t_i aisitei-ru
         everyone-ACC someone-NOM   love
         'Someone loves everyone' ('every > some' is OK.)

     b. Dareka -o$_i$ daremo -ga t_i aisitei-ru
         someone-ACC everyone-NOM   love
         'Everyone loves someone' ('every > some' is OK.)

(7) is in the basic SOV order and the subject takes wide scope over the object. In (8), the object is scrambled over the subject. In this case, the scrambled object can take scope over the subject as in (8a), and more importantly, the subject can take scope over the object as shown in (8b). Thus, scope ambiguity obtains with movement.

Given this generalization, examples such as (9) provide a clue for the structure of existential sentences.

(9) Itutu-izyoo -no   ike -ni san -syurui-izyoo -no   sakana-ga i-ru
      five-more than-GEN pond-in three-kind -more than-GEN fish-NOM be
      'There are more than three kinds of fish in every pond'
      (more than five > more than three, *more than three > more than five)

This example has the Locative-Nominative-V order. The locative can take wide scope over the nominative phrase. But the opposite scope relation is impossible.

Suppose that there are 15 ponds altogether. The sentence is false under the situation where four kinds of fish are found in more than five of them but none of the ponds has more than two kinds of fish. This is expected if the sentence is in basic word order. Then, it should be unambiguous like (7). On the other hand, if the sentence is derived from the Nominative-Locative-V order by preposing the locative phrase, it is incorrectly predicted to be ambiguous. Hence, Kuno concludes that the basic word order of existential sentences is Locative-Nominative-V.

Although Kuno does not examine the scope relation of nominative phrases and negation, his discussion suggests that nominative phrases in existential sentences are generated quite low in the structure and take fairly narrow scope, that is, below the locative. And in fact, nominative phrases in existential sentences fall under the scope of negation, as shown in (10).

(10) Kono ike -ni-wa sakana-ga   i-na-i
      this   pond-in-TOP fish   -NOM be-not
      'There is no fish in this pond' (not > some)

This sentence does not mean that there is a fish such that it is not in this pond. It seems then that nominative phrases take scope under negation, contrary to Kozumi's claim.

This problem, however, is only apparent. First, if we substitute *koi-dake* 'carp-only' for *sakana* 'fish' in this example as in (11), the nominative phrase does take wide scope over negation.

(11) Kono ike -ni-wa koi -dake-ga   i-na-i
      this   pond-in-TOP carp-only-NOM be-not
      'It is only carp that this pond does not have' (only > not)

The situation is a familiar one. Lasnik (1999) considers the following contrast with quantifier lowering:

(12) a. Some politician$_i$ is likely [t_i to address John's constituency]
         (some > likely, likely > some)

    b. Every coin$_i$ is 3% likely [t_i to land heads]
         (every > 3% likely, *3% likely > every)

As discussed extensively in May (1977), examples like (12a) allow the narrow scope reading of the raised subject. However, Lasnik points out that this phenomenon is limited to those NPs that can be construed as indefinites, and does not obtain with quantified NPs such as *every coin*, as shown in (12b). He suggests then that there is no quantifier reconstruction with A-movement, and

speculates that the narrow scope reading of a raised indefinite is possible for an independent reason.

A precise account for the contrast in (12) is pursued in Lasnik (2008). However, let us simply assume for the purpose here that reconstruction in A-chains (or more generally Case chains) is difficult with quantified NPs but is readily available or even forced with indefinites. Then, the contrast between (10) and (11) is precisely what is expected under Koizumi's approach. In both examples, the nominative phrase originates within VP and is raised to a position in the T-projection so that its Case is licensed. As a result, it c-commands negation. But in the case of (10), the raised nominative phrase takes reconstructed scope below negation because it is an indefinite.

This line of analysis predicts that a contrast similar to the one in (10) and (11) should obtain with nominative objects. Koizumi's important observation was that a nominative quantified NP in object position takes scope over negation. If this is because the NP moves to the T-projection for Case licensing as Koizumi argues, then an indefinite nominative object should be able to take reconstructed scope under negation. The prediction is borne out by the examples in (13) and (14).

(13) a. Kiyomi-wa Taroo-dake-o   hihan -deki-na-i
     -TOP -only-ACC criticize-can -not
   'Kiyomi cannot criticize only Taroo' (not > can > only)

  b. Kiyomi-wa Taroo-dake-ga   hihan -deki-na-i
      -TOP -only-NOM criticize-can -not
    'It is only Taroo that Kiyomi cannot criticize' (only > not > can)

(14) Kiyomi-wa hito -o /-ga   hihan -deki-na-i
     -TOP person-ACC/-NOM criticize-can -not
  'Kiyomi cannot criticize a person' (not > can > only)
  (Not 'there is someone who Kiyomi cannot criticize')

(13) confirms that *Taroo-dake* 'only Taroo' takes scope over negation when it is a nominative object. (14), on the other hand, shows that an indefinite object takes scope under negation whether it is in nominative or accusative. This example, like (10), is a counter-example to the generalization that nominative phrases take scope over negation. But it is consistent with Koizumi's movement analysis of nominative objects, provided that indefinites take reconstructed scope.

The following contrast, discussed in detail in Miyagawa (2003), seems to pose another problem for Koizumi's analysis:

(15) a. Zen'in-ga   sono tesuto-o   uke -na -katta (yo /to omo-u)
         all   -NOM that test -ACC take-not-Past   Part that think
         'All didn't take that exam' (all > not, *not > all)

   b. Sono tesuto-o$_i$ zen'in-ga t_i uke -na -katta (yo /to omo-u)
         that test -ACC all   -NOM   take-not-Past   Part that think
         'That exam, all didn't take' (all > not, not > All)

(15a) has SOV order, and the quantified NP in the subject position takes scope over negation. However, when the object is scrambled over the subject, the latter exhibits scope ambiguity with negation as in (15b). This already indicates that quantified NPs with nominative Case do not necessarily take scope over negation.

   Miyagawa's analysis of this contrast is as follows. The subject in (15a) is in TP Spec, satisfying the EPP requirement of T, and hence, takes scope over negation, as illustrated in (16a).

(16) a. [$_{TP}$ subject$_i$ [$_{T'}$ [$_{NegP}$ [$_{Neg'}$ [$_{vP}$ t_i [$_{v'}$ [$_{VP}$ object V] v]] Neg]] T]]

   b. [$_{TP}$ object$_i$ [$_{T'}$ [$_{NegP}$ [$_{Neg'}$ [$_{vP}$ subject [$_{v'}$ [$_{VP}$ t_i V] v]] Neg]] T]]

(15b), on the other hand, can be derived in two ways. First, it can be derived from (16a) by adjoining the object to TP by A'-scrambling. In this case, the subject takes scope over negation. Second, the object can be scrambled to TP Spec and consequently satisfy the EPP requirement of T instead of the subject as in (16b). When this happens, the subject remains in vP Spec and as a result, takes narrow scope under negation.

   I slightly modified this analysis in Saito (2009) on the basis of examples like (17).

(17) a. Zen'in-ga   zibun-zisin-ni   toohyoosi-na -katta (to omo-u)
         all   -NOM self -self -DAT vote   -not-Past   that think
         'Everyone did not vote for herself/himself' (all > not, *not > all)

   b. Zibun-zisin-ni$_i$ zen'in-ga t_i toohyoosi-na -katta (to omo-u)
         self -self -DAT all   -NOM   vote   -not-Past   that think
         'For herself/himself, everyone did not vote' (all > not, not > all)

(17) shows the same kind of contrast as (15). In particular, the scrambling in (17b) makes the narrow scope construal of the subject possible. However, the scrambled dative phrase cannot be in TP Spec because that would make the example a Condition (C) violation. It follows that in this example also the subject is in TP Spec and satisfies the EPP requirement of T. I argued then that the

element in TP Spec can scopally interact with negation, but the sentence-initial constituent, whether it is a subject or a scrambled object, is in the Spec position of a higher functional head, which I called Pred, and takes scope over negation. According to this analysis, the structures of (15a-b) are as in (18a-b) respectively.

(18) a. [$_{PredP}$ subject$_i$ [$_{Pred'}$ [$_{TP}$ t_i' [$_{T'}$ [$_{NegP}$ [$_{Neg'}$ [$_{vP}$ t_i [$_{v'}$ [$_{VP}$ object V] v]] Neg]] T]] Pred]]

b. [$_{PredP}$ object$_i$ [$_{Pred'}$ [$_{TP}$ subject$_j$ [$_{T'}$ [$_{NegP}$ [$_{Neg'}$ [$_{vP}$ t_j [$_{v'}$ [$_{VP}$ t_i V] v]] Neg]] T]] Pred]]

This analysis assumes crucially that quantified NPs in TP Spec can scopally interact with sentential negation. It seems then to be in contradiction with Koizumi's, which attributes the wide scope property of nominative objects to their movement to TP Spec. But here again, the problem is only apparent. Note first that if we substitute *Taroo-dake* 'Taroo-only' for *zen'in* 'all' in (15), the subject takes scope over negation, as shown in (19).

(19) a. Taroo-dake-ga   sono tesuto-o   uke -na -katta (yo /to   omo-u)
           -only-NOM that test -ACC take-not-Past   Part that think
           'Only Taroo didn't take that exam' (only > not, *not > only)

b. Sono tesuto-o$_i$ Taroo-dake-ga t_i uke -na -katta (yo /to   omo-u)
     that test -ACC   -only-NOM   take-not-Past   Part that think
     'That exam, only Taroo didn't take' (only > not, *not > only)

Secondly, when *zen'in* 'all' appears as a nominative object, it scopally interacts with sentential negation.[3]

(20) a. Kiyomi-wa   zen'in-o   hihan -deki-na -katta
           -TOP all   -ACC criticize-can -not-Past
           'Kiyomi couldn't criticize everyone' (not > all)

b. Kiyomi-wa   zen'in-ga   hihan -deki-na -katta
           -TOP all   -NOM criticize-can -not-Past
           'Kiyomi couldn't criticize anyone/everyone' (not > all, all > not)

Hence, the contrast between Koizumi's examples with NP-*dake-ga* 'NP-only-NOM' and Miyagawa's (15b) with *zen'in-ga* 'all-NOM' should be attributed to the lexical properties of the quantified NPs. Both move to TP Spec. The former takes wide scope over negation in this position while the latter can be construed with wide or narrow scope. This receives indirect support from the English examples in (21).

---

[3] The verb 'criticize' consists of two morphemes *hihan* 'criticism' and *su* 'do'. The latter has an irregular potential form *deki* 'can do'.

(21) a. Everyone didn't take that exam (every > not, not > every)
b. Most people didn't take that exam (most > not)
c. Only John didn't take that exam (only > not)

(21a) is ambiguous. *Everyone* is clearly in TP Spec, and yet, it can take wide or narrow scope with respect to negation. On the other hand, it is known that some quantified NPs take wide scope over negation when they are in TP Spec. *Most people* in (21b) is a typical example. (21c) shows that *only John* exhibits this pattern as well. Then, *zen'in* 'all' in Japanese behaves like *everyone*, and *Taroo-dake* 'only Taroo' is like *only John*, which is hardly surprising.

The generalization that nominative phrases take scope over sentential negation, which Koizumi's analysis seems to imply, cannot be maintained. In particular, indefinites take narrow scope and *zen'in* 'all' can take wide or narrow scope with respect to negation. However, I argued in this section that the relevant examples support Koizumi's movement analysis of nominative objects. All nominative phrases are licensed within the T-projection. But indefinites take reconstructed scope and *zen'in* can scopally interact with negation when they are in TP Spec. Thus, the complex pattern of the scope of nominative phrases follows from Koizumi's analysis.

3. Covert A'-movement for Case Licensing

In the preceding section, I presented a supporting argument for Koizumi's hypothesis that nominative objects move to a position in the T-projection for the purpose of Case licensing. Koizumi argues that this movement is overt A-movement to the inner Spec of TP. In this section, I will provide evidence that it is covert A'-movement. I will first show that the movement does not obey the locality conditions on A-movement. Then, I will suggest that the movement is covert.

Recall Koizumi's hypothesis that nominative objects move to the inner Spec of TP. The illustration in (6) is repeated below in (22).

(22) [$_{TP}$ Kiyomi [$_{TA}$[$_{T'}$ [$_{NegP}$ [$_{VP}$ [$_{VP}$ right eye only-NOM close] -can] -Neg] -Pres.]]]

One problem arises immediately with this analysis. Since nominative objects undergo A-movement to TP Spec, we would expect them to exhibit subject properties, but they do not. The subject-oriented anaphor *zibun* 'self' is often used as a diagnosis for subjecthood. As shown in (23), only subjects qualify as antecedents of *zibun*.

Mamoru Saito

(23) a. Hanako$_i$-ga   Taroo$_j$-o   zibun$_{i/*j}$-no ie   -de sikat-ta (koto)
                  -NOM   -ACC self   -GEN house-in scolded fact
       '(the fact that) Hanako scolds Taroo in her house'

    b. Taroo$_i$-ga Hanako$_j$-niyotte zibun$_{i/*j}$-no ie   -de t_i siakar-are-ta (koto)
                  -NOM   -by   self   -GEN house-in   scolded-was   fact
       '(the fact that) Taroo was scolded by Hanako in his house'

(23a) indicates that *zibun* cannot take an object as its antecedent, and the passive sentence in (23b) shows that the relevant notion of 'subject' is surface subject. And it is well known that nominative objects do not qualify as subjects in this regard. (24) confirms this generalization.

(24) Hanako$_i$-ga   Taroo$_j$-ga   zibun$_{i/*j}$-no   ie   -de sikar -e-ru (koto)
                  -NOM   -NOM self   -GEN house-in scold-can   fact
        '(the fact that) Hanako can scold Taroo in her house'

A more direct problem for Koizumi's analysis concerns the locality of the movement in question, as is also pointed out in Takahashi (2008). As noted in Inoue (1976), there is an important difference between passive and nominative object constructions. To illustrate this, let us first consider the example of causative in (25).

(25) a. Hanako$_i$-wa Taroo$_j$-ni   (zibun$_{i, j}$-no   heya -de) hon -o   kak -ase-ta
                  -TOP   -DAT self   -GEN room-in book-ACC write-made
        'Hanako makes Taroo study linguistics in her/his room'

    b. Hanako$_i$-wa [$_{vP}$ Taroo$_j$-ni [$_{VP}$ (zibun$_{i, j}$-no heya -de) hon-o kak]]-ase-ta

(25a) looks like a simple sentence with a complex verb *kak-ase* 'write-make', but both the agent *Hanako* and the causee *Taroo* are possible antecedents for *zibun*. It has thus been widely assumed since Kuroda (1965) that causatives involve sentential embedding with the causee as the embedded subject. I assume, following Murasugi and Hashimoto (2004), among others, that the causative verb -*(s)ase* takes a *v*P complement as in (25b). This analysis straightforwardly accounts for the contrast in (26) with passive.

(26) a. Taroo-ga$_i$ Hanako-niyotte [$_{vP}$ t_i [$_{VP}$ hon -o   kak]]-ase-rare-ta (koto)
                  -NOM   -by   book-ACC write-made-was   fact
        '(the fact that) Taroo was made by Hanako to write a book'

    b. *Hon -ga$_i$ Hanako-niyotte [$_{vP}$ Taroo-ni   [$_{VP}$ t_i kak]]-ase-rare-ta (koto)
        book-NOM   -by   -DAT   write-made-was   fact
        '*Lit.* (the fact that) a book was made by Hanako for Taroo to write'

In (26a), the causee is moved to the matrix subject position. In (26b), on the other hand, the object is raised across the embedded *v*P Spec. The result is ungrammatical as expected since the movement is a minimality violation.[4] What is interesting is that the object in (25a) cannot be passivized as shown in (26b) but can be turned into a nominative object with the addition of the suffix *-(rar)e* 'can'. This is shown in (27).

(27)   Hanako-ga [_{*v*P} Taroo-ni [_{VP} hon -ga   kak]]-ase -rare-ru (koto)
               -TOP   -DAT   book-NOM write-make-can   fact
       '(the fact that) Hanako can make Taroo write a book'

The grammaticality of this example is totally unexpected if the nominative object undergoes A-movement to a position in the matrix T-projection. The movement should then take place across the embedded *v*P Spec and should violate minimality exactly like (26b). Based on evidence of this kind, I concluded in Saito (1982) that nominative objects are not licensed by A-movement to TP Spec.

    As we saw in the proceeding section, the scope properties of nominative objects indicate that they move to a position within the T-projection. At the same time, (27) shows that the movement is not to an A-position. Then, we are led to the conclusion that the relevant movement is A'-movement. This conclusion, which is also argued for in Takahashi (2008), solves the puzzle mentioned at the outset of this section as well. Recall that nominative objects are not possible antecedents for *zibun*, and thus, lack subject properties. This is expected if they move to an A'-position in the T-projection. If T can host multiple Specs and TP Specs can be A or A' positions, it seems plausible that the A'-Specs are at the outer edge, as illustrated in (28).

(28)   [_{TP} A' [_{T'} subject [_{T'} A [_{T'} [_{*v*P} …] T]]]]

I will henceforth assume that nominative objects move to the outer Spec of TP.
    The next issue is whether the movement in question is overt or covert. First, Yatsushiro (1999) presents arguments that nominative objects stay within VP in overt syntax. One of her arguments is based on examples of VP-preposing such as (29).

(29) a. Kai-ga   eigo   -ga   yom-e -sae si-ta (no -ni -wa odoroi-ta)
            -NOM English-NOM read-can-even did   that-DAT-TOP surprised
         '(I was surprised that) Kai was even able to read English'

---

[4]   There is an issue regarding why the NP-movement cannot proceed through the edge of *v*P. A proposal that excludes this derivation will be made in Section 4 below.

b. [$_{VP}$ Eigo-ga yom-e-sae]$_i$ [$_{TP}$ Kai-ga t_i si-ta] (no-ni-wa odori-ta)

c. *[$_{VP}$ Yom-e-sae]$_i$ [$_{TP}$ Kai-ga eigo-ga t_i si-ta] (no-ni-wa odori-ta)

Japanese has an interesting construction, exemplified by (29a), where VP is followed by a focus particle and is embedded under the main verb *su* 'do'. In this construction, the embedded VP can be preposed by scrambling. Yatsushiro points out that a nominative object must be preposed with the VP in this case, as shown in (29b-c). This is expected if it is within VP but not if it moves to the Spec of TP in overt syntax.

Although I believe that Yatsushiro's argument is valid, I also noticed that some consider (29b) marginal. This is not surprising as VP-preposing is best with agentive verbs.[5] But once it is assumed that nominative objects move to the outer Spec of TP, another argument can be constructed to show that the movement is not overt. One of the reasons that Koizumi assumed that the landing site is the inner Spec of TP is because SOV seems to be the basic word order even in nominative object sentences. This intuition can be confirmed by the distribution of floating quantifiers (FQs). It is well known that Japanese employs floating quantifiers extensively as in (30).

(30) a. Gakusee-ga   san-nin   sake-o   non-da
         student -NOM 3   -person sake-ACC drank
         'Three students drank sake'

   b. Gakusee-ga   sake-o   san-bon   non-da
       student -NOM sake-ACC 3   -bottle drank
       'A student drank three bottles of sake'

As noted in Kuroda (1980) and Haig (1980), an FQ must occur adjacent to the associate NP and hence (31a) is not grammatical.[6]

(31) a. ??Gakusee-ga   sake-o   san-nin   non-da
          student -NOM sake-ACC 3   -person drank
          'Three students drank sake'

   b. Sake-o$_i$ gakusee-ga t_i san-bon   non-da
       sake -ACC student -NOM   3   -bottle drank
       'A student drank three bottles of sake'

(31b) is an apparent counter-example to this generalization as the subject intervenes between the object and an FQ associated with it. Kuroda and Haig

---

[5] See Hoshi (1995) and Saito (2006) for detailed discussion of VP-preposing in Japanese.

[6] More precisely, they must be in sister relation as discussed in detail in Miyagawa (1989).

take this as evidence that OSV order is derived by scrambling: the object originates in a position adjacent to the FQ and is preposed to the sentence-initial position. This analysis assumes that the ungrammatical (31a) has the structure in (32a).

(32) a. [$_{TP}$ subject [$_{VP}$ object FQ$_{subj}$ V]]

b. [subject$_i$ [object$_j$ [$_{TP}$ t_i FQ$_{subj}$ [$_{VP}$ t_j V]]]]

In particular, it presupposes that the example cannot have the structure in (32b), where both the object and the subject are scrambled to the sentence-initial position. If (32b) were possible, the example would be incorrectly predicted to be grammatical.[7]

Given this, let us consider the example in (33) with a nominative object, which is perfectly grammatical in SOV order.

(33) a. Hanako-ga   terugugo-ga   hanas-e-ru
           -NOM Telugu -NOM speak-can
    'Hanako can speak Telugu'

b. [subject$_i$ [$_{TP}$ Nom object$_j$ [$_{T'}$ t_i (FQ$_{subj}$) [$_{VP}$ t_j V]]]]

If the nominative object moves overtly to the outer Spec of TP, the subject must be scrambled to its left as in (33b) so that the SOV order is obtained. Then, it is predicted that an FQ associated with the subject can occur after the nominative object as indicated. But this prediction fails. (34b) has the ungrammatical status equivalent to (31a) while (34a) is perfect.

(34) a. Gakusee-ga   san-nin   terugugo-ga   hanas-e-ru
      student -NOM 3 -person Telugu -NOM speak-can
      'Three students can speak Telugu'

b. ??Gakusee-ga   terugugo-ga   san-nin   hanas-e-ru
      student -NOM Telugu -NOM 3 -person speak-can
      'Three students can speak Telugu'

Thus, the ungrammaticality of (34b) shows that a nominative object does not move to the outer Spec of TP overtly, and that the movement has to be covert.

---

[7] In order to exclude (33b), I suggested in Saito (1985) that subjects cannot be scrambled. Comrie (1987), on the other hand, appeals to the ban on crossing. More recently, Ko (2007) and Takita (2008) proposed an account based on Fox and Pesetsky's (2005) theory of linearization. The last account seems most principled, but it is unfortunately incompatible with the analysis of short scrambling proposed later in this paper.

Mamoru Saito

4. EPP, Case Licensing, and VP-internal Scrambling

So far, I argued that nominative objects move covertly to the outer Spec of TP to have the nominative Case licensed.[8] This means that nominative Case licensing need not takes place in overt syntax. And this in turn implies that the EPP, which requires TP Spec to be filled overtly, is independent of nominative Case licensing. In a simple example like (35), T attracts the nominative subject to its Spec to satisfy the EPP requirement.

(35)   [$_{TP}$ Taroo-ga$_i$ [$_{vP}$ t_i [$_{VP}$ kantongo -o   hanas-u]]] (koto)
                -NOM   Cantonese-ACC speak   (fact)
      '(the fact that) Taroo speaks Cantonese'

As a result, the nominative phrase happens to move overtly to a position where its Case is licensed. But this is accidental, and the movement does not take place for the purpose of nominative licensing.

Ura (1999), who entertains a covert feature movement analysis for the Case licensing of nominative objects, also reaches this conclusion. He argues that T attracts the closest NP to its Spec regardless of the Case. For example, it has been known that 'dative subjects' exhibit subject properties despite the fact that they are not nominative. This is illustrated in (36) with the subject-oriented anaphor *zibun*.

(36)   Hanako$_i$-ga /-ni   Taroo$_j$-ga   zibun$_{i/*j}$-no ie   -de sikar -e-ru (koto)
             -NOM/-DAT   -NOM self   -GEN house-in scold-can fact
      '(the fact that) Hanako can scold Taroo in her house'

Ura proposes that dative subjects are assigned inherent Case in *v*P Spec and are attracted by the EPP-feature of T to TP Spec.

The hypothesis that the EPP-feature of T attracts the closest NP raises an interesting question in relation with scrambling. As shown in (37b), a direct object can be scrambled to a sentence-medial position following the subject in TP Spec.

(37) a. Hanako-ga   Taroo-ni   hon -o   okut-ta (koto)
                -NOM   -DAT book-ACC sent   fact
      '(the fact that) Hanako sent a book to Taroo'

   b. Hanako-ga   hon -o$_i$ Taroo-ni t_i okut-ta (koto)
              -NOM book-ACC   -DAT   sent   fact

---

[8]   Takahashi (2008) also argues that nominative objects take wide scope due to covert A'-movement, but proposes that the relevant movement is QR. The discussion in this section is consistent with this alternative.

If this kind of short scrambling can be to the edge, i.e., the outer Spec, of vP, why is it that an object cannot be attracted to TP Spec instead of the subject as in (38)?

(38) $[_{TP}$ object$_i$ $[_{vP}$ t_i' $[_{v'}$ subject $[_{VP}$ t_i V]]]]

Once the object is scrambled to the outer Spec of vP, the EPP-feature of T would automatically attract it as it is the closest NP. The nominative Case on the subject can be licensed if the subject raises covertly to the outer Spec of TP exactly like a nominative object.

As noted above, Miyagawa (2003) argues that an object can move to TP Spec as in (38) and check the EPP-feature of T instead of the subject, but we have seen evidence that it is always the subject that moves to TP Spec and satisfies the EPP requirement. Further, a scrambled object does not have subject properties. For example, it can never be the antecedent of *zibun*, as shown in (39).

(39) Taroo-o$_j$ Hanako$_i$-ga t_j zibun$_{i, *j}$-no   heya -de sikar-u (koto)
        -ACC   -NOM   self   -GEN room-in scold   fact
    '(the fact that) Hanako scolds Taroo in her room'

Thus, the derivation in (39) must somehow be excluded.

Here, I would like to propose that the outer Spec of vP in (38) is necessarily an A'-position, and consequently, the movement of the object to TP Spec is ruled out as an instance of improper movement from an A'-position to an A-position. I hypothesized above that the outer Spec of TP is an A'-position, as shown in (40).

(40) a. $[_{TP}$ A' $[_{TP}$ subject $[_{T'}$ …   (=(28))

    b. $[_{vP}$ A' $[_{vP}$ subject $[_{v'}$ A […

This can be stated more precisely as follows. Let us assume, following Chomsky (1995), that the uninterpretable EPP-feature of T must be checked as soon as it enters the derivation. In (40a), the subject checks this feature and at this point, all the lexical requirements of T are satisfied. We may then say that any Spec created after this point counts as an A'-position. If we extend this reasoning to vP Specs, then any Spec external to the subject must be an A'-position. This is so because v needs to discharge its theta-role to the subject, but once this is done, it has no further lexical requirements to be satisfied. Then, a vP Spec external to the subject is an A'-position while an internal vP Spec is an A-position, as illustrated in (40b).

Given this, when an object is scrambled to the outer Spec of vP, the derivation necessarily crashes. The EPP-feature of T attracts it to TP Spec

because it is the closest NP but this results in improper movement. The only possible landing site for scrambling to the edge of *v*P, then, is the inner Spec. That is, the scrambling takes place, and then, the subject is merged to receive the external theta-role. In this case, the EPP-feature of T attracts the subject as it is the closest NP.

Although this may sound somewhat speculative, it serves to solve one outstanding problem with the analysis of scrambling. While scrambling across the subject exhibits both A and A' properties, scrambling to the edge of *v*P shows strict A properties, as discussed in detail by Mahajan (1990), Tada (1993) and Nemoto (1993), among others. The examples in (41) demonstrate that an object preposed to the sentence-initial position can serve as the antecedent for an anaphor.

(41) a. ?*[Otagai -no sensee]-ga Masao-ni karera-o   syookaisi-ta (koto)
          each other-GEN teacher-NOM   -DAT they -ACC introduced   fact
         'Lit. Each other's teachers introduced them to Masao'

b. [Karera-o$_i$ [$_{TP}$ [otagai-no sensee]-ga Masao-ni t_i syookaisi-ta]] (koto)

Scrambling of this kind patterns with A-movement in this respect. On the other hand, scrambling of an anaphor to the sentence-initial position does not induce a Condition (C) violation, as shown in (42).

(42) a. [Taroo-to Hanako]-ga Masao-ni otagai   -o   suisensi-ta   (koto)
               -and   -NOM   -DAT each other-ACC recommended fact
         '(the fact that) Taroo and Hanako recommended each other to Masao'

b. [Otagai-o$_i$ [$_{TP}$ [Taroo-to Hanako]-ga Masao-ni t_i suisensi-ta]] (koto)

Here, scrambling exhibits an A'-property. The use of *A* and *A'* is quite confusing in the literature on scrambling, but let us assume for concreteness, following Webelhuth (1992), that scrambling across the subject is non-operator A'-movement. The scrambled object in (41b) is a possible antecedent for the anaphor because it is in a non-operator position, and (42b) is grammatical because the object is in an A'-position.

Scrambling to the post-subject position shows a different pattern. First, it shares the non-operator property with scrambling across the subject. (43) illustrates this.

(43) a. ?*Masao-ga [otagai -no sensee]-ni karera-o   syookaisi-ta (koto)
               -NOM each other-GEN teacher-DAT they -ACC introduced   fact
         'Masao introduced them to each other's teachers'

b. Masao-ga [karera-o$_i$ [[otagai-no sensee]-ni t_i syookaisi-ta]] (koto)

c. [Karera-o$_i$ [$_{TP}$ Masao-ga [$_{vP}$ t_i' [[otagai-no sensee]-ni t_i syookaisi-ta]]]] (koto)

(43b) shows that scrambling of *karera* 'they' to the post-subject position enables the NP to serve as the antecedent for an anaphor contained in the indirect object. But scrambling to the post-subject position exhibits a paradigm different from (42) when the scrambled NP is an anaphor. The scrambling of *otagai* 'each other' over its antecedent results in ungrammaticality as in (44).

(44) a. Masao-ga [Taroo-to Hanako]-ni   otagai -o   suisensi-ta   (koto)
         -NOM   -and   -DAT each other-ACC recommended fact
      '(the fact that) Masao recommended each other to Taroo and Hanako'

   b. *Masao-ga [otagai-o$_i$ [[Taroo-to Hanako]-ni t_i suisensi-ta]] (koto)

   c. *[Otagai-o$_i$ [$_{TP}$ Masao-ga [$_{vP}$ t_i' [[Taroo-to Hanako]-ni t_i suisensi-ta]]]] (koto)

(44c) is particularly interesting. There is only one necessary intermediate landing site: the movement must proceed through the edge of the *v*P phase. The final landing site should not cause a Condition (C) violation because it is a (non-operator) A'-position as we saw in (42b). Then, the intermediate landing site must be an A-position.

    It has been a mystery why the vP Spec position of the intermediate trace in (44c) counts as an A-position. But this is precisely what we expect given the discussion in this paper. Recall the structure of *v*P edge in (40b), repeated below in (45).

(45)   [$_{vP}$ A' [$_{vP}$ subject [$_{v'}$ A […

If the scrambled object moves to the outer Spec, it is in A'-position. But then, the derivation crashes because the EPP-feature of T attracts this NP, forcing an improper movement. Hence, the derivation converges only when the landing site of scrambling is the inner Spec, which is an A-position. The subject, then, moves to TP Spec to satisfy the EPP requirement of T, and the object undergoes further scrambling to the sentence-initial position in the case of (44c). Thus, the paradigm in (44) receives an account based on Ura's (1999) hypothesis that the EPP attracts the closest NP, independently of nominative Case licensing.

5. Concluding Remarks

In this paper, I argued that nominative objects move covertly to the outer Spec of TP in order to have the nominative Case licensed. In the course of the discussion, I provided evidence for Koizumi's (1995, 1998) movement analysis, comparing the scope properties of quantifiers and indefinites. At the same time, I argued that the movement is covert and its landing site is an A'-position. Finally, I discussed

Mamoru Saito

a consequence of the analysis, i.e., that the EPP-feature attracts the closest NP to its Spec independently of nominative Case licensing. I argued that this hypothesis, proposed by Ura (1999), enables us to account for the strict A-property of scrambling to the post-subject position.

There are other consequences that need to be examined. It should be noted first that an issue remains with the precise identity of the covert A'-movement of nominative objects. I argued in this paper that it takes place for nominative Case licensing, extending Tada's (1992) and Koizumi's (1995, 1998) analyses. On the other hand, as noted in Fn.8, it is argued by Takahashi (2008) that it is QR. In either case, the movement applies to satisfy a requirement of the moved item. But if the Case licensing analysis is correct, it provides stronger support for Bošković (2007), which argues for the last resort nature of movement in general. According to this analysis, the movement does not take place to eliminate any deficiency of T, but only to license the nominative Case on the object. Hence, it implies that there are movements for Case licensing that are motivated solely by the moved item.

Another consequence concerns the analysis of scope rigidity of quantifiers. As noted above, Japanese exhibits scope rigidity as in (7), repeated below in (46).

(46) Dareka -ga   daremo -o   aisitei-ru
      someone-NOM everyone-ACC love
      'Someone loves everyone' (some > every, *every > some)

This generalization holds even with nominative objects, as (48) shows.

(47) Dareka -ga   daremo -ga   nagur-e-ru
      someone-NOM everyone-NOM hit   -can
      'Someone can hit everyone' (some > every, *every > some)

Given the analysis proposed in this paper, the object in (47) moves covertly to the outer Spec of TP, creating the configuration in (48).

(48) [$_{TP}$ daremo-ga$_i$ [$_{T'}$ dareka-ga$_j$ [[$_{vP}$ t_j [[$_{VP}$ t_i V]] v] T]]]

This seems problematic as *daremo* 'everyone' occupies a position higher than *dareka* 'someone'. Takahashi (2008) takes this as evidence that the scope relation between two quantified NPs is determined by their surface order. He argues in particular that examples like (47) constitute evidence for Bobaljik and Wurmbrand's (2008) ScoT, a soft constraint whose effects are most visible in scrambling languages such as German and Japanese.

(49)　Scope Transparency (ScoT)
　　　If the order of two elements at LF are A > B, the order at PF is A > B.

An alternative would be to appeal to the Pred projection alluded to in Section 2. Recall the hypothesis that the sentence-initial constituent is in PredP Spec in Japanese. Given this, the subject, *dareka-ga* 'someone-NOM', in (47) occupies this position in overt syntax. Thus, after the covert movement of the object, *daremo-ga* 'everyone-NOM', to the outer Spec of TP, the configuration in (50) obtains.

(50) [$_{PredP}$ dareka-ga$_j$ [[$_{TP}$ daremo-ga$_i$ [$_{T'}$ t_j' [[$_{vP}$ t_j [[$_{VP}$ t_i V]] v] T]]] Pred]]

This is consistent with the rigidity effect in (40), but it remains to be seen whether this account extends to all instances of scope rigidity observed in Japanese.

References

Bošković, Željko. 2007. On the locality and motivation of move and agree: An even more minimal theory. *Linguistic Inquiry* 38: 589-644.

Bobaljik, Jonathan, and Susi Wurmbrand. 2008. Word order and scope: Transparent interfaces and the 3/4 signature. Ms., University of Connecticut.

Chomsky, Noam. 1995. *The minimalist program*. Cambridge, Mass.: MIT Press.

Comrie, Akiko K. 1987. On so-called quantifier floating in Japanese. Ph.D. dissertation, USC.

Fox, Danny, and David Pesetsky. 2005. Cyclic linearization of syntactic structure. *Theoretical Linguistics* 31: 1-45.

Haig, John. 1980. Some observations on quantifier floating in Japanese. *Linguistics* 18: 1065-1083.

Hoshi, Hiroto. 1994. Passive verbs and light verbs: A study on theta-role assignment. Ph.D. dissertation, University of Connecticut.

Inoue, Kazuko. 1976. *Henkeebunpoo-to nihongo [Transformational grammar and the Japanese language]*. Tokyo: Taishukan.

Ko, Heejeong. 2007. Asymmetries in scrambling and cyclic linearization. *Linguistic Inquiry* 38: 49-83.

Koizumi, Masatoshi. 1995. Phrase Structure in Minimalist Syntax. Ph.D. dissertation, MIT.

Koizumi, Masatoshi. 1998. Remarks on Nominative Objects. *Journal of Japanese Linguistics* 16: 39-66.

Mamoru Saito

Kuno, Susumu. 1973. *The structure of the Japanese language*. Cambridge, Mass.: MIT Press.

Kuroda, S.-Y. 1965. Causative forms in Japanese. *Foundations of Language* 1: 30-50.

Kuroda, S.-Y. 1971. Remarks on the notion of subject with reference to words like *also*, *even*, and *only*, Part II. *Annual Bulletin* 4: 127-152, Research Institute of Logopedics and Phoniatrics, University of Tokyo.

Kuroda, S.-Y. 1980. Bonkoozoo-no hikaku [Comparison of sentence structures]. In *Nichieigo hikaku kooza 2: Bunpoo [Lectures on comparative studies of English and Japanese 2: Grammar]*, ed. Tetsuya Kunihiro, 23-61. Tokyo: Taishukan.

Lasnik, Howard. 1999. Chains of arguments. In *Working minimalism*, eds. Samuel D. Epstein and Norbert Hornstein, 189-215. Cambridge, Mass.: MIT Press.

Lasnik, Howard. 2008. Lecture notes for Advanced Syntax, Nanzan University.

Mahajan, Anoop K. 1990. The A/A-bar distinction and movement theory. Ph.D. disseration, MIT.

May, Robert. 1977. The grammar of quantification. Ph.D. dissertation, MIT.

Miyagawa, Shigeru. 1989. *Structure and Case marking in Japanese.* San Diego: Academic Press.

Miyagawa, Shigeru. 2003. A-movement scrambling and options without optionality. In *Word order and scrambling*, ed. Simin Karimi, 177-200. Oxford: Blackwell.

Murasugi, Keiko, and Tomoko Hashimoto. 2004. Three pieces of acquisition evidence for the *v*-VP frame. *Nanzan Linguistics* 1: 1-19.

Nemoto, Naoko. 1993. Chains and Case positions: A study from scrambling in Japanese. Ph.D. dissertation, University of Connecticut.

Nomura, Masashi. 2005. Nominative Case and agree(ment). Ph.D. dissertation, University of Connecticut.

Saito, Mamoru. 1982. Case marking in Japanese: A preliminary study. Ms., MIT.

Saito, Mamoru. 1985. Some asymmetries in Japanese and their theoretical implications. Ph.D. dissertation, MIT.

Saito, Mamoru. 2006. Expletive replacement reconsidered: Evidence from expletive verbs in Japanese. In *Form, Structure, and Grammar*, eds. Patrick Brandt and Eric Fuss, 255-273. Berlin: Akademie Verlag.

Saito, Mamoru. 2009. Optional A-scrambling. *Japanese/Korean Linguistics* 16: 44-63.

Sano, Masaki. 1985. LF movement in Japanese. *Descriptive and Applied Linguistics* 18: 245-259.

Tada, Hiroaki. 1992. Nominative objects in Japanese. *Journal of Japanese Linguistics* 14: 91-108.

Tada, Hiroaki. 1993. A/A-bar partition in derivation. Ph.D. dissertation, MIT.

Takahashi, Masahiko. 2008. Case-valuation, phasehood, and nominative/accusative conversion in Japanese. Presented at NELS 39, Cornell University.

Takita, Kensuke. 2008. Remarks on the proper binding condition effects on overt and covert merge. Ms., Nanzan University.

Webelhuth, Gert. 1992. *Principles and Parameters of Syntactic Saturation*. Oxford: Oxford University Press.

Ura, Hiroyuki. 1999. Checking theory and dative subject constructions in Japanese and Korean. *Journal of East Asian Linguistics* 8: 223-254.

Yatsushiro, Kazuko. 1999. Case licensing and VP structure. Ph.D. dissertation, University of Connecticut.