Deriving Greenberg’s Universal 20 and Its Exceptions

Guglielmo Cinque

Of the 24 mathematically possible orders of the four elements demonstrative, numeral, adjective, and noun, only 14 appear to be attested in the languages of the world. Some of these are unexpected under Greenberg’s Universal 20. Here it is proposed that the actually attested orders, and none of the unattested ones, are derivable from a single, universal, order of Merge (Dem > Num > Adj > N) and from independent conditions on phrasal movement.

Keywords: DP, word order, nominal modifiers, typology

1 A Derivational Account of Greenberg’s Universal 20

Greenberg’s (1963) Universal 20,\(^1\) under its most sensible interpretation (see Hawkins 1983: 117ff.), states (a) that in prenominal position the order of demonstrative, numeral, and adjective (or any subset thereof) conforms to the order Dem > Num > A, and (b) that in postnominal position the order of the same elements (or any subset thereof) conforms either to the order Dem > Num > A or to the order A > Num > Dem.

Forty years later, the first part of this statement remains (virtually) unchallenged,\(^2\) while the second part has proven both too restrictive and too permissive. Some studies, for example, have uncovered the existence of postnominal orders that are excluded by Greenberg’s formulation. For

I thank audiences at UCLA, New York University, and the Universities of Padua and Venice, where parts of this material were presented in 2003 and 2004—in particular, Laura Brugè, Giuliana Giusti, K. A. Jayaseelan, Ed Keenan, Hilda Koopman, Anoop Mahajan, Cecilia Polletto, Walter Schweikert, Dominique Sportiche, Michal Starke, Tim Stowell, and Anna Szabolcsi. I also thank Paola Benincà, Richard Kayne, Jan Rijkhoff, and two anonymous reviewers for their helpful comments on a previous draft of the article. Special thanks go to Dominique Sportiche, who encouraged me to explore the implications of the analysis in a systematic fashion.

In line with much work in linguistic typology and generative grammar, I am assuming that all languages have demonstratives, numerals, and adjectives as adnominal modifiers in their DP. This is by no means an innocent assumption, especially given the often-made claim that some languages have no distinct class of adjectives (or numerals), but use in their stead NPs (in PPs or appositions) or verbs (in relative clauses). See, for example, Schachter 1985:13ff. and Rijkhoff 2002:327ff. I would, nonetheless, like to keep to that assumption, thinking of Baker 2003:chap. 4 and Dixon 2004. Also see the introduction to Cinque, to appear.

Num refers throughout to cardinal numerals, not to Number (singular, plural, etc.) or to ordinal numerals.

\(^1\) ‘‘When any or all of the items (demonstrative, numeral, and descriptive adjective) precede the noun, they are always found in that order. If they follow, the order is either the same or its exact opposite’’ (p. 87).

\(^2\) See, for example, Lu 1998:165: ‘‘[H]ead-final NPs have D Q A N [= Dem Num A N] as the only basic order.’’ Croft and Deligianni (2001:7) state that ‘‘[i]f more than one modifier occurs before the noun, the order is always Dem < Num < Adj; the only exception to this generalization is an alternative word order in a language that otherwise conforms to the generalization (Alamblak), and other possible orders in Korean and Quechuan. If more than one modifier occurs after the noun, however, virtually any order appears to be possible as the basic order of modifiers.’’ The alternative order Croft and Deligianni refer to is Dem A Num N, which is also reported as an alternative order in Mongolian by Whitman (1981:414). This apparent exception should be considered with caution, however. The four languages are OV and all have prenominal relative clauses, which characteristically appear between Dem and Num (or to the left of Dem). See Cinque 2003b and, for Alamblak, Bruce 1984:106ff. So this alternative order might well involve adjectives in a reduced relative clause rather than in attributive position (see Whitman 1981). The same analysis may extend to the special
Gabra (Cushitic), Luo (Nilotic), and Logoli (Bantu), Heine (1981) reports the order N Num A Dem, which conforms neither to Dem > Num > A nor to A > Num > Dem. Hawkins (1983: 119), citing Hyman (1979:27), mentions the existence in Aghem (Bantu) of the order N A Dem Num, which again conforms neither to Dem > Num > A nor to A > Num > Dem; and he reports, also citing Hyman (1981:31), that Noni (Bantu), in addition to N Dem Num A, displays the order N Dem A Num, again unexpected under Greenberg’s formulation.

On the basis of these facts, Hawkins (1983:119–120) concludes that the second part of Greenberg’s Universal 20 must be abandoned, and that for the postnominal order of demonstrative, numeral, and adjective essentially every combination is possible.³ His proposed revision of Greenberg’s Universal 20 is quoted in (1).

(1) ‘‘When any or all of the modifiers (demonstrative, numeral, and descriptive adjective) precede the noun, they (i.e., those that do precede) are always found in that order. For those that follow, no predictions are made, though the most frequent order is the mirror-image of the order for preceding modifiers. In no case does the adjective precede the head when the demonstrative or numeral follow.’’ (Hawkins 1983:119–120, (20′))

Given that certain postnominal order possibilities (namely, *N Num Dem A and *Num N Dem A)⁴ are (still) unattested, as far as I was able to determine (see below), and given that even the actually attested orders differ significantly, as we will see, in the percentage of languages that instantiate them, I will not follow Hawkins, or Croft and Deligianni, in their conclusion that postnominally anything goes; instead, I will propose a refinement of an analysis I suggested in Cinque 1996, 2000 to derive Greenberg’s basic generalization—one that may also derive its exceptions, and the different degree of markedness of the various orders.

The analysis I suggested in those works aimed at deriving the essential left-right asymmetry in word order possibilities found prenominally (one) and postnominally ((at least) two), starting from the idea (actually forced by Kayne’s (1994) Antisymmetry Theory) that generating modifiers symmetrically to the left and to the right of the N could not easily account for the absence, prenominally, of the order A Num Dem.⁵ This asymmetry could instead be made sense of, I

³ As is apparent from the quotation in footnote 2, this conclusion is shared by Croft and Deligianni (2001).

⁴ While the orders of Gabra, Luo, Logoli, Aghem, and Noni mentioned above show that Greenberg’s original formulation is too restrictive (in forbidding their existence), the fact that Num N Dem A (which Greenberg’s formulation admits) is unattested shows that his formulation may also be too permissive. While it is not impossible that such an order will be documented in some yet-to-be-studied language, I will submit below that it is not attested for principled reasons. Hawkins (1983:118–119) claims that two other orders permitted by Greenberg’s formulation (namely, Dem A Num and Num A N Dem) are unattested in his expanded sample of 336 languages. Such orders, however, have since been documented. See (6).

⁵ For example, one could assume the two following symmetric base structures, the right-branching (ia) and the left-branching (ib):
submitted, if all orders are derived by moving (or not moving) the NP around the modifiers, base-generated prenominally in the fixed order Dem Num A.

If nothing moves, the unique (Merge) order found prenominally (the Dem Num A N order) surfaces. As for the two postnominal orders, they arise via the two ways in which the NP raises: either alone, from specifier to specifier (Spec to Spec) of Agr(eement) projections found above each of the functional projections hosting adjectives, numerals, and demonstratives, to give the order N Dem Num A (as in (2)), or by moving successively to each such Spec and pied-piping the category that dominates it, in a ‘‘roll-up’’ fashion that reverses the order of the modifiers, to give N A Num Dem (as in (3)).

(2)

(i) a. [Dem [Num [A [N]]]]

This would account for the mirror-image order of Dem Num A found to the right of the N. In order to account for the other postnominal order (the same as that found prenominally: N Dem Num A), one would have to either assume the left-branching structure [[[N] Dem] Num] A] (possibly at odds with the natural relative semantic scope of these elements) or admit the possibility that N(P) raises to the left of Dem in the base-generated structure (ia). In either case, though, it would not be clear how the unwanted order A Num Dem N could be excluded, as nothing principled in such a system would seem to prevent the symmetric right-branching structure [A [Num [Dem [N]]]] once [[[N] Dem] Num] A] is allowed, or, alternatively, the raising of N(P) to the right of Dem in the base-generated structure (ib). To exclude them, specific ad hoc principles would have to be introduced. The principled unavailability of such symmetric solutions is precisely one of the main consequences of Antisymmetry Theory (see Kayne 1994).

6 On their possible status as Agr projections, see footnote 24, based on evidence discussed in Shlonsky 2004:sec. 6. Nothing crucial would change if demonstratives and numerals were to be heads rather than maximal projections in Spec position (a question that remains to be clarified). In Cinque 1996, 2000, to derive the order N Dem Num A, I actually posited N-movement, though the same order could be derived by moving the NP from Spec to Spec (without pied-piping). Here, because of such redundancy, and, more crucially, because N-movement will prove unable to exclude the unattested orders, only phrasal movement (of NP—or of a larger XP including NP) will be assumed to be available.
This was a simplification in that that analysis generated only the orders in (4) and, taking partial movements into consideration, those in (5)—6 out of the 24 orders that are the mathematically possible combinations of the four elements Dem Num A N (factorial $4 = 4 \times 3 \times 2 \times 1 = 24$).

(4) a. Dem Num A N
   b. N Dem Num A
   c. N A Num Dem

(5) a. Dem Num N A
   b. Dem N Num A
   c. Dem N A Num

The crucial question then is, of the 24 orders, which ones are actually attested? And, if more are attested than the 6 indicated, how can the attested ones be derived in this system without also deriving the unattested ones?

(6) shows all the 24 orders. The ‘‘√’’ and ‘‘*’’ preceding them indicate whether the order is attested or unattested, respectively. This indication is based on the typological (or other) sources available in the literature on the order of N, demonstrative, numeral, and adjective (that I have been able to find). The ‘‘Ø’’ and references following some of the orders indicate that in those references, the order in question is explicitly claimed not to be attested. (Orders enclosed in a box are by far the most common. See the remark in Hawkins’s revised formulation of Greenberg’s
Universal 20 in (1) about the mirror image of the prenominal order being the most frequent order in postnominal position. This is indeed matched by the relatively few languages instantiating the order N Dem Num A, as already noted in Greenberg’s remark quoted in footnote 10 below.}

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>✓</td>
<td>Dem</td>
<td>Num</td>
<td>A</td>
</tr>
<tr>
<td>b.</td>
<td>✓</td>
<td>Dem</td>
<td>Num</td>
<td>N</td>
</tr>
<tr>
<td>c.</td>
<td>✓</td>
<td>Dem</td>
<td>N</td>
<td>Num</td>
</tr>
<tr>
<td>d.</td>
<td>✓</td>
<td>N</td>
<td>Dem</td>
<td>Num</td>
</tr>
<tr>
<td>g.</td>
<td>*</td>
<td>Num</td>
<td>N</td>
<td>Dem</td>
</tr>
<tr>
<td>k.</td>
<td>✓</td>
<td>A</td>
<td>N</td>
<td>Dem</td>
</tr>
<tr>
<td>l.</td>
<td>✓</td>
<td>N</td>
<td>A</td>
<td>Dem</td>
</tr>
</tbody>
</table>


\(^8\) According to Rijkhoff (1998:357), “[t]he order [Dem Num N A] is . . . rather frequent in Europe.” Outside Europe it is documented in, among other languages, Yao (Jones 1970), Burushaski, Guarani (Rijkhoff 2002:328), Abkhaz, Farsi, Kiowa, Mam (Croft and Deligianni 2001), Cape Verdean, Mauritian, and Seychelles Creoles, Kristang, Kriyol, and Tok Pisin (Haddican 2002).

\(^9\) This order is documented in Sampur and Camus (Heine 1981) (but see Rijkhoff 2002:274–275), and in Maasai (Koopman 2003). According to Croft and Deligianni (2001:7), it is also a possible alternative order (of the Dem N A Num order) in Hualapai and Lahu.

\(^10\) Greenberg (1963:87) states that N Dem Num A is “[a] less popular alternative” to N A Num Dem, citing Kikuyu as one example. Other languages that apparently display this order are Turkana, Rendille (Heine 1981), Noni (Hyman 1981:31), Nkore-Kigu (Lu 1998:162n59, 165), Abu’ (Lynch 1998:171), Arbore (Hayward 1984:212), Bai and Moro (Dryer 2003:20, 43). It also appears as a possible alternative order in Romanian (Corniles cu 1992:212); but see Cinque 2004 for discussion.

\(^11\) This order is found in Koiari (which also has the order N A Dem Num with most adjectives; Dutton 1996:60ff.) and in Bai (Wiersma 2003:669); according to Dryer (2000:20), Bai also has N Dem Num A as an alternative order. [A N]-def Num is also an alternative order of the unmarked Dem Num A N order of Icelandic (Sigurðsson 1993:194, Vangsnes 2004). The possibility of this order in Koiari and Bai (and of the order A N Dem Num in Gude and Sango; see below) indicates that the last sentence of Hawkins’s (1983) revision of Greenberg’s (1963) Universal 20, given in (1), may be too strong. Greenberg’s Universal 18 was less categorical (“When the descriptive adjective precedes the noun, the demonstrative, and the numeral, with overwhelmingly more than chance frequency, do likewise”). This is because of the existence, noted by Greenberg, of “a small number of instances (e.g., Efik) in which the demonstrative follows while the adjective precedes” (p. 86). Greenberg (1989:113) also cites Zande. See also Dryer 2000:34.

\(^12\) This order is found in Lalo (Björnerud 1998:116ff.), Lisu (Bradley 2003:228–229), Akha (Hansson 2003:241), Qiang (LaPolla 2003b:39), Aghem (Hyman 1979:27), Port Sandwich (Crowley 2002:653), Koiari (Dutton 1996:60ff.), which also has the order A N Dem Num with certain adjectives, and Lingala (Haddican 2002). Croft and Deligianni (2001) also assign Babungo and, more tentatively, Woleaian to this order.
Keeping to the idea that no symmetric base-generation of modifiers is possible, and that postnominal orders are only a function of the raising of the NP (or of an XP containing the NP),


14 Among the languages that instantiate this order are Kabardian and Warao (Hawkins 1983:119), Burmese, Lolo, Maru, Rawang (Jones 1970), Ladakhi (Koshal 1979:108), Gambian Mandinka (Rijkhoff 1998:356), Cuna (Quesada 1999:232), Kaki Ae (Clifton 1995:46), and Pech (Holt 1999:62ff.). It is an alternative order for N A Num Dem in Kunama (Bender 1996:41), and for Dem N Num A in Hualapai and Lahù (Croft and Deligianni 2001:7).

15 According to Hawkins (1983:119) and Lu (1998:165), this order is not attested. However, Rijkhoff (2002:328) and Haddican (2002) report Berbice Dutch Creole as instantiating it. Haddican (2002) also documents this order for the creole languages Sranan and Bislama. Lynch (2002:769–770, 781, 809) gives it as the order of Xäräcùù, Iaai, and Puluwatese. To judge from Siewierska and Uhlirová 1997:132–133, Polish and Russian also have this order as an alternative order to Dem Num A N.

16 According to Heine (1981) and Chiao (1998), this order is not attested. However, as noted, Lu (1998:162), Harriehausen 1990:144), in Wolof (Sy 2003), and in a number of creoles (Haddican 2002). It is also displayed by the Australian language Watjarri (Douglas 1981:241).

17 According to Lu (1998:162), this order is not attested. However, as noted, Heine (1981) documents it in three languages: Gabra, Logoli, and Luo (on Luo, also see Chiao 1998), Noonan (1992:154) documents it in Lango. Ross (2003:132) and Tryon (2002:576) give it as the order of Kele and Buma, respectively. Croft and Deligianni (2001) give it as an alternative order of Manam.

18 According to Hawkins (1983:119) and Lu (1998:165), this order is not attested. However, Thornell (1997:71) and Haddican (2002) give it as the order of Sango, and Rijkhoff (1998:356, 358, 2002:332n19) mentions (doubtfully) the possible existence of two other languages with this order: Gude and Zande.

19 This order is found in Cambodian, Javanese, Karen, Khmu, Palaung, Shan, Thai (Rijkhoff 1990:32), Enga (Lynch 1998:171), Dagaare (Bodomo 1993), Ewe (Essegbey 1993), Gungbe (Aboh 1996, 2004), Labu and Ponapean (Lynch 1998:121), Mao Naga (Girdhar 1994:452), Selepet, Yoruba (Hawkins 1983:119), West Greenlandic (which also has N A Dem Num as an alternative order) (Rijkhoff 2002:326), Amele, Igbo, Kusaeian, Manam (Croft and Deligianni 2001), Fa d’Ambu, Nubi (Haddican 2002), Kugu Nganhcara (Smith and Johnson 2000:388), Cabècar (Quesada 1999:232), Kunama (Bender 1996:41), and Māori (Pearce 2002).
it seems possible to derive all the attested orders, without also deriving the unattested ones. What
we have to assume is this:

(7) a. Merge order: \([\ldots [\text{WP Dem} \ldots [\text{XP Num} \ldots [\text{YP A [[NP N]]]]]\ldots]]\)

b. Parameters of movement:
   i. No movement (unmarked), or
   ii. Movement of NP plus pied-piping of the *whose picture* type\(^{21}\) (unmarked), or
   iii. Movement of NP without pied-piping (marked), or
   iv. Movement of NP plus pied-piping of the *picture of who* type\(^{22}\) (more marked
      still).
   v. Total (unmarked) versus partial (marked) movement of NP with or without pied-
      piping (in other words, NP raises all the way up, as in (6d,l,p,t,x), or just partially, as in
      (6b,c,k,n,o,r,s,w), around its modifiers).
   vi. Neither head movement nor movement of a phrase not containing the (overt)
      NP is possible (except perhaps for focus-related movements of phrases to a DP-
      initial position).\(^{23}\)

Let’s consider how these assumptions manage to derive the attested orders, and fail to derive
the unattested ones (in the computation of markedness, I take the markedness induced by partial
movement to be less severe than the markedness induced by movement without pied-piping,
which is in turn less severe than that induced by movement with pied-piping of the *picture of
who* type; see below for some discussion).

(6a) (Dem Num A N) is derived if nothing moves (7bi). (No marked option: very many
languages.)

(6b) (Dem Num N A) is derived from Dem Num A N if NP raises one notch, around A,\(^{24}\)
either with (vacuous) pied-piping of the *whose picture* type (7bii) (unmarked) or without pied-

---

\(^{20}\) Although I assume Move (when it takes place) to be interspersed with Merge, in accord with the Extension
Condition (Chomsky 1995:327–328), in (7a) I have just indicated the applications of Merge, for simplicity.

\(^{21}\) Namely, movement of [NP [XP]].

\(^{22}\) Namely, movement of [XP [NP]]. This option corresponds to Shlonsky’s (2004) notion of freezing.

\(^{23}\) On the possible, marked preposing of APS to DP-initial position (often for focusing purposes), see among others

One additional parameter is the *obligatory versus optional* application of movement. For example, the alternative
orders Q Dem Num N A, Q Dem N A Num, Q N A Num Dem, N A Num Dem Q of Standard Arabic (see Fassi Fehri
1999, Cinque 2000, Shlonsky 2004) point to the obligatory character of movement of the NP around the adjectives
followed by optional movement (plus pied-piping of the *whose picture* type) around numerals, demonstratives, and
universal quantifiers.

Partly on the basis of data from Jones 1970, Nguyen (2004:chap. 4) also proposes to derive the different orders
of various Southeast Asian languages (the Dem Num Classifier N A of Yao; the Dem N A Num Cl of Burmese, Lolo,
Maru, Lahu, Râwang and ancient Chinese; the Num Cl N A Dem of Vietnamese, Malay, Indonesian, Nung, Hmong,
White and Black Tai, Sedang, Stre, Katu, Cham, and Brōu; and the N A Num Cl Dem of Thai, Khmer, Lao, Javanese,
Khmor, Shan, Palaung, and Karen) “from the same underlying structure,” Dem Num Cl N (A), in terms of two movement

\(^{24}\) As anticipated above, I assume the NP to raise to the Spec of a functional projection merged in between YP and
Num of (7a), with similar functional projections available in the position of the other dots of (7a). (See below for further
discussion.) In trees (2) and (3), these functional projections were labeled AgrP, as some actual agreement process takes
place, in some languages, when the NP (with or without pied-piped material) raises to their Spec. For example, Shlonsky
piping (7biii) (marked). (Despite the markedness of partial movement, it includes the unmarked case of pied-piping: many languages.)

(6c) (Dem N Num A) is derived if NP moves two notches, around A and Num (i.e., partially—marked option) without pied-piping (7biii)—marked option). (Two marked options: very few languages.)

(6d) (N Dem Num A) is derived if NP moves three notches, around A, Num, and Dem (i.e., all the way up) without pied-piping (7biii): marked). (One marked option: few languages.)

(6e) (Num Dem A N) cannot be derived through (7). NP has not moved, and the modifiers to its left are in the wrong Merge order (cf. (7a)).

(6f) (Num Dem N A) cannot be derived through (7). Raising of NP without pied-piping implies a wrong Merge order of the modifiers (Num Dem A N) (see (7a)). Raising of NP with pied-piping of the picture of who type either of [Dem N] or of [Num Dem N] also implies a wrong Merge order (either Num A [Dem N] or A [Num Dem N]).

(6g) (Num N Dem A) cannot be derived through (7). Raising of NP without pied-piping implies that the Merge order is Num Dem A N, which is a wrong order. Raising of NP with pied-piping of the whose picture type again implies a wrong Merge order of the modifiers (Num A Dem N), with N first raising around Dem and [N Dem] then raising around A. Raising of NP with pied-piping of the picture of who type (raising of [Num N] two notches) also implies a wrong Merge order of the modifiers (Dem A Num N).25

(6h) (N Num Dem A) cannot be derived through (7). Raising of NP without pied-piping implies a wrong Merge order (Num Dem A N). Raising of NP with successive pied-pipings of the whose picture type also implies a wrong Merge order (A Dem Num N). Raising of NP without pied-piping around Dem and Num, followed by raising with pied-piping around A, would derive (6h), but, again, from a wrong Merge order (A Num Dem N). (Similarly if NP were to move

(2004:secs. 5, 6) discusses cases of demonstrative agreement in Moroccan Arabic (had lə wlad ‘this the boys’ vs. lə wlad hadu ‘the boys these’) and of cardinal numeral agreement (in postnominal, though not in prenominal, position) in the Syrian dialect spoken in Palmyre.

25 But see the problematic case of Michif mentioned in footnote 2.

26 There is another potential derivation, from a Merge order like A Num Dem N, with NP raising around Dem and [N Dem] subsequently pied-piping Num ([Num [N Dem]]) around A. But this is again a wrong Merge order of the modifiers.

Potentially problematic is the case of Kilivila. Senft (1986:105) gives Num N Dem A as the general order for this language. However, he also gives the order Dem Num A N as possible:

If a speaker of Kilivila wants to refer to

‘these two beautiful girls’.

he has to indicate the noun class of ‘girls’ with the classifier for ‘female’ in the demonstrative pronoun, in the numeral, and in the adjective. Thus this sentence runs in Kilivila:

minasina nayu namanabweta vivila

which translates literally as

mi-na-si-na na-yu na-manabweta vivila

this-female-PL-this female-two female-beautiful girls

In this example the classificatory particle na is underlined . . . (p. 69)

The first order (Num N Dem A) would only cease to be a problem if the adjective there derived from a relative clause, which is higher than Num, with Num N moved past the relative clause and Dem. See footnote 2.
around Num and pied-pipe it to the left of A and then move on without further pied-pipings. The Merge order in this case would be Dem A Num N—again, the wrong order.)

(6i) (A Dem Num N) cannot be derived through (7). NP has not moved, and the modifiers to its left are in the wrong Merge order (see (7a)).

(6j) (A Dem N Num) cannot be derived through (7). NP has moved one notch, but the two modifiers to its left are in the wrong Merge order (see (7a)). (6j) could also arise via raising of NP with pied-piping of the picture of who type of either Dem N or A Dem N around Num, but both derivations presuppose a wrong Merge order (A Num Dem N and Num A Dem N, respectively).

(6k) (A N Dem Num) has a well-formed, though marked, derivation with raising of NP plus pied-piping of the picture of who type of the lowest modifier (A), followed by raising of [A N] without pied-piping around both Num and Dem. (Two marked options: very few languages.)

(6l) (N A Dem Num) has a derivation in which NP raises past A, followed by pied-piping of the whose picture type past Num, followed by raising of [N A] without pied-piping (marked) past Dem. (One marked option: few languages.)

(6m) (Dem A Num N) cannot be derived through (7). NP has not moved, and the modifiers to its left are in the wrong Merge order (see (7a)). (See footnote 2 for discussion of the apparent existence of some such cases.)

(6n) (Dem A N Num) has a derivation with partial (marked) raising of NP plus pied-piping of the picture of who type of [A N] (marked) around Num. (Two marked options: very few languages.)

(6o) (Dem N A Num) has a derivation from (7a) involving partial (marked) raising of NP plus pied-piping of the whose picture type, vacuously, and nonvacuously (of [N A]) around Num. (One marked option: many languages.)

(6p) (N Dem A Num), if genuine (see footnote 27), may be especially marked, as its derivation from (7a) would seem to involve raising of NP with successive pied-pipings of the whose picture type around A and Num (alternatively, a single raising of the picture of who type of [A N] around Num) and then extraction of the sole NP around Dem.27

(6q) (Num A Dem N) cannot be derived through (7). NP has not moved, and the modifiers to its left are in the wrong Merge order (see (7a)).

(6r) (Num A N Dem) has a derivation with partial (marked) raising of NP plus pied-piping of the picture of who type of A and Num ([Num A N]) (marked) around Dem. (Two marked options: very few languages.)

(6s) (Num N A Dem) has a derivation with partial (marked) raising of NP around A, followed by raising plus pied-piping of the picture of who type of [Num N A] (marked) around Dem. (Two marked options: few languages (but see footnote 32).)

27 The literature known to me reports only three languages with this order: Pijiantijatjara (Bowe 1990:111); Noni, which has (6d) (N Dem Num A) as its primary order (Rijkhoff 2002:273); and Nkore-Kiga (Dryer 2003:43), which also has (6d) as an alternative order (Lu 1998:162n59, 165). It is possible, thinking of its prevailing status as an alternative order, that this order is actually spurious (with A a reduced relative clause; see footnote 2) and that such subextractions should be ruled out entirely.
(6t) (N Num A Dem) has a derivation with raising of NP without pied-piping around A and Num (marked), followed by raising plus pied-piping of the *whose picture* type of [N Num A] around Dem. (One marked option: few languages.)

(6u) (A Num Dem N) cannot be derived through (7). NP has not moved, and the modifiers to its left are in the wrong Merge order (see (7a)).

(6v) (A Num N Dem) cannot be derived through (7). Raising of NP without pied-piping implies a wrong Merge order of the modifiers (A Num Dem N) (see (7a)). Raising of NP with pied-piping of the *picture of who* type either of [Num N] or of [A Num N] also implies a wrong Merge order (either A Dem [Num N] or Dem [A Num N]).

(6w) (A N Num Dem) has a derivation from (7a) with raising of NP plus pied-piping of the *picture of who* type of A around Num (marked), followed by raising of [A N Num] around Dem. (One marked option: few languages.)

(6x) (N A Num Dem) has a derivation from (7a) involving raising of NP with successive pied-pippings of the *whose picture* type all the way up. (No marked option: very many languages.)

The fact that all N-final orders that do not respect the order Dem Num A ((6e), Num Dem A N; (6i), A Dem Num N; (6m), Dem A Num N; (6u), A Num Dem N) are very clearly unattested can indeed be taken to indicate that it is the raising of NP (or of an XP containing it) that is responsible for word order variation within the DP (perhaps, more generally, that it is the raising of the lexical part of a phrase that is responsible for word order variation within its ‘extended projection’).

---

28 Although its derivation involves one marked option, this order is in fact manifested by very few languages. One of three languages instantiating it (Gude) has, as an alternative order, (6k). The two orders differ minimally with respect to derivation, it seems, in that the raising of [A N] around Num pied-pipes Num around Dem in (6w) but fails to do so in (6k). Quite generally, the study of alternative orders within a single language should provide interesting insights into the question of which parametric choices cooccur. See, for example, the orders in Koiari mentioned in footnotes 11 and 12 (A N Dem Num (with a few adjectives) and N A Dem Num (with most adjectives)), a situation that recalls Romance and Celtic, if we ignore the further raising around Dem and Num. The orders of Bai (N Dem Num A and A N Dem Num (footnotes 10 and 11)) appear to involve either raising of NP without pied-piping or raising of NP and A with pied-piping of the *picture of who* type. Also see the alternative orders of Hualapai and Lahu (Dem N Num A and Dem N A Num (footnote 9), differing in terms of pied-piping); those of Kunama (Dem N A Num and N A Num Dem (footnote 14), differing in terms of partial vs. total “roll-up”); those of West Greenlandic (N A Num Dem and N A Dem Num (footnote 19); those of Bangla and Syrian Arabic (Dem Num A N and Dem A N Num (footnote 13)); those of Manam (N Num A Dem and N A Num Dem (footnotes 17 and 19)); and those of Polish and Russian (Dem Num A Num and Num A Num Dem (footnote 15)). Polish, apparently alone among the Slavic languages, also has the order Dem Num N A (only with classificatory adjectives). See Sussex 1975, Willim 2000, and references cited there.

29 If the universal Merge order is . . . Num Cl . . . N, as suggested in Den Dikken 2003:6–7 (see also Nguyen 2004: chap. 4 and Simpson 2005), the fact that Thai and other Southeast Asian languages have N A Num Cl Dem order with the adjectives in the opposite of the Chinese/English order (Sproat and Shih 1991, Den Dikken and Singhapreecha 2004) suggests that the NP, which moves via pied-piping of the *whose picture* type, must skip the head between Num and Cl. According to Albro (1998:3), Nawdm, which has N A Num Dem order with the demonstrative ‘this’, and Dem N A Num with the demonstrative ‘that’, has the adjectives in the English order, not the reverse. This would seem to imply movement of NP around the As without pied-piping followed by movement plus pied-piping around Num (and Dem, in the case of ‘this’) (see Albro 1998).

30 Allowing for movement of both N and NP, and especially for remnant movement of phrases not containing the NP (or containing only its trace), would wrongly permit the derivation of most of the unattested orders. See below for a constraint on remnant movement, proposed by Kayne (2005), that actually appears to ban such possibilities.
This offers a way to make sense of the fact that only to the right of N are more orders possible (indeed, those deriving from the different modes in which NP, or an XP containing it, raises). It also offers a way to derive, at least in part, the different degrees of markedness of each order (and, I take it, the ensuing differences in the numbers of languages that instantiate each one).

Although I know of no clear independent reason why movement without pied-piping should count as more marked than movement with pied-piping (of the whose picture type) (whence the respective numbers of languages instantiating each order), it seems natural that those orders whose derivation crucially involves pied-piping of the whose picture type should be less marked (and be instantiated in more languages) than those involving pied-piping of the picture of who type. The different degree of markedness of the two types of movement appears to be suggested independently by contrasts like the following in English (and corresponding contrasts in other languages): Whose pictures did you see yesterday? versus ?Pictures of who did you see yesterday?; Now I know whose picture he saw yesterday versus *Now I know a picture of who he saw yesterday.

2 Questions and Implications

The account sketched above raises a number of questions, which given our present state of knowledge can only receive very speculative answers. Here, I will briefly consider the following:

(8) a. What triggers the movement of the NP (with or without pied-piped material) within the DP?
   b. Why are there languages that do not have movement of the NP?
   c. Where does the NP (with or without pied-piped material) move to?
   d. Why is movement of phrases other than the NP (with or without pied-piped material) unavailable?

One possible answer to (8a) and (8b) relates to the presumable need for the various phrases that make up the “extended” projection of the NP (in Grimshaw’s (1991) sense) to be licensed. Suppose that each phrase (the one containing an Adjective Phrase, the one containing the Number Phrase, the one containing the Demonstrative Phrase, etc.) needs to be endowed with a nominal feature to be licensed (i.e., to count as part of the extended projection of NP), and that this can

31 The orders derived with pied-piping of the whose picture type ((6b), (6o), and (6x)) are represented by many languages, as opposed to the orders derived without pied-piping ((6c) and (6d)). A possible basis for the markedness of movement of the NP without pied-piping is discussed below in connection with the trigger of movement of the NP.
32 Indeed, it appears that the orders that involve pied-piping of the picture of who type ((6k), (6n), (6r), and (6w)) are instantiated in very few languages, fewer than those derived without pied-piping (i.e., (6c) and (6d)) or those derived only partially by pied-piping of the whose picture type (i.e., (6l)).

A possible exception may be represented by (6s), which apparently incorporates pied-piping of the picture of who type as one step of the derivation and yet seems to be instantiated by a fair number of languages (see footnote 16).

As noted, partial movements seem to be more marked than ones sending the NP all the way up (whence perhaps the smaller number of languages that instantiate the order in (6c) as opposed to that in (6d), and that in (6b) as opposed to that in (6a)).
be brought about by merging above it an Agr(eement) head whose Spec ultimately comes to have such a nominal feature, either by movement of NP, or by Merge of such a feature, which enters in an agreement relation with the NP without movement (the Agree operation of Chomsky 2000).

Some languages employ the first mode, others employ the second mode, and still others employ both, with the result that some will have movement all the way up, some will have no movement, and others will have partial movement of the NP. This also provides an answer to (8c). The further (sub)question why languages, in the unmarked case, move NP together with pied-piped material (of the whose picture type) is far less clear. Here, I will very tentatively submit that it may have to do with a general condition on movement/atraction proposed by Kayne (2005:sec. 5.6), who suggests that what moves to the Spec of a functional head H is the category closest to H that is not the complement of H (nor, I will add, the specifier of the complement of H).

The crucial question is how ‘‘closest to H’’ is defined. Consider (9).

(9) The category closest to H is the category c-commanded by H that is dominated by the fewest number of nodes (where ‘‘node’’ includes every node, whether ‘‘category’’ or ‘‘segment,’’ in Kayne’s (1994) sense).

For example, in a structure like (10) (derived by moving NP to Spec.Agr2P), this definition singles out Agr2P as the category (distinct from its complement) closest to Agr1. It is c-commanded by Agr1 and is dominated by fewer nodes than either NP, ZP, or AP.

\[ \ldots [_{\text{Agr1P}} \text{Agr1} [_{\text{YP NumP}} [_{\text{YP Y}} [_{\text{Agr1P[NP N]}} [_{\text{Agr2P}} \text{Agr2} [_{\text{ZP AP \ldots [NP N]}]}]]]]] \]

I take (9) to be the unmarked definition of ‘‘closest to H,’’ with the result that only Agr2P, not [NP N] alone, will raise to Spec.Agr1, thus deriving the pied-piping option (of the whose picture type) as the only unmarked option. For [NP N] in (10) to count as ‘‘closest to’’ Agr1, we would have to modify the definition in (9), introducing a limitation on the type of node that counts in the calculation of ‘‘closest to H’’: namely, ‘‘category’’ only, rather than the more general ‘‘node’’ (which includes both ‘‘categories’’ and ‘‘segments’’). If we do this, the higher Agr2P ‘‘segment’’ in (10) no longer counts, so that Agr2P and [NP N] will be equally close to Agr1.33

Let us consider now the last question in (8). We have noted that certain phrases other than NP (with pied-piped material)—namely, specifiers (APs)—must be allowed to move up alone (under special and limited conditions); see footnote 23. Yet no movement of any other sort must be permitted if we do not want the prenominal orders that we ruled out with the assumptions in (7) to be derivable. In particular, ‘‘head’’ movement and ‘‘remnant’’ movement will have to be unavailable. If we could move N, or NP, beyond AP, as in (10), and then move the remnant ZP across NumP (and DemP), we would end up with the orders A Dem Num N and Dem A Num

33 Strictly speaking, to capture the fact that the (few) languages that move the NP alone do not also avail themselves of the pied-piping option, one would also have to introduce a clause stating that [NP N] counts as closer to Agr1 than Agr2P (this might be achieved by exploiting the fact that [NP N] c-commands, and is not c-commanded by, Agr2P according to Kayne’s (1994) definition of c-command).
N, which are not found as the exclusive (or unmarked) order in any language. However, this result is already achieved under the assumptions just sketched. Putting aside the special focus position, there is no other landing site for the “remnant” than Spec,Agr₁ (or the Spec of a higher Agr). But Kayne’s condition—that what moves to the Spec of a functional head H is the category closest to H (that is not the complement of H)—will ensure that Agr₂P, not the “remnant” ZP, will be attracted to Spec,Agr₁.

3 Further Questions and Implications

Two of the crucial components of the analysis suggested above are that [Dem . . . [Num . . . [A . . . [N]]]] is the (universal) structure of the DP resulting from Merge, and that word order variations within DP across languages are fundamentally a function of how (different phrases containing) the NP move up the structure. At least, such assumptions were seen to be able to derive the actually attested orders of the four elements crosslinguistically (without also deriving the unattested ones). Other plausible options, such as raising just the N, and/or moving phrases not including the NP, were seen not to yield the same results.

Should this general approach to crosslinguistic word order variation within DP be confirmed, it will be natural to ask whether the quite extensive crosslinguistic word order variation within the clause and other phrases should not also be treated in terms of movement of “extended” phrases (necessarily) containing the “lexical” projection (VP, AP, PP, etc.), rather than in terms of head movement. This question remains to be explored.

[Dem . . . [Num . . . [A . . . [N]]]] is but a fragment of the internal structure of the DP. If we were to add universal quantifiers, ordinals, numeral classifiers, and relative clauses (RCs)—setting aside the fact that A is just an abbreviation for an ordered sequence of adjectives (Cinque 1994, Scott 2002, and references cited there), and ignoring Case, Number, possessors, demonstrative reinforcing, various types of determiners, functional adjectives like other and same (Kayne 2005: sec. 2.1), diminutives/augmentatives, complements,³⁴ and so on—we would have 8 elements, whose mathematically possible combinations number (factorial 8 = ) 40,320.

The actually possible combinations would (luckily) be much fewer if the 8 elements entered a fixed hierarchical structure resulting from Merge, and if variations of this structure could only arise via upward movements of phrases containing the NP.

³⁴ Complements require at least a brief discussion. They do not seem to be part of the NP that raises. So, for example, in Semitic, except for construct state genitives, which are found between the initial N and its modifiers, prepositional complements are stranded at the end of the DP, and are not dragged along by the NP in its “roll-up” movements (see Cinque 2000, Shlonsky 2004, and references cited there). Similarly, in Romance, the NP can raise across (certain classes of) adjectives, stranding its PP complements (see Cinque 2003a). This nonadjacency of complements follows from Kayne’s analysis of the (overt and covert) prepositions that introduce them. In Kayne 2000, 2002, 2004, prepositions are argued to be heads merged higher up in the extended projection of the NP (or outside the DP altogether), attracting their “complements” and forcing (in VO languages) the remnant to raise to their left, thus making them final in the DP. Interestingly, complements of the N in OV languages are generally DP initial, before Dem. This is the case in, for example, Turkish (Jaklin Kornfilt, pers. comm.); Hindi (Anoop Mahajan (pers. comm.) notes that, more markedly, they can also occur after the N); and Malayalam (though, as K. A. Jayaseelan (pers. comm.) points out, they are necessarily introduced as predicates of relative clauses). In other words, they seem to involve attraction to the left of P, but no movement of the remnant.
There is suggestive evidence that universal quantifiers are located higher than Dem;\(^{35}\) ordinal numerals between Dem and cardinal ones (see Shlonsky 2004); numeral classifiers between Num and A (see Den Dikken 2003, Simpson 2005; also see Kayne 2005, for evidence that languages that do not have overt classifiers may have covert ones); and RCs between Dem/ordinals and cardinals, though reduced RCs are possibly (also) located below cardinals (see Cinque 2003b). As a first approximation, this would give (11),\(^{36}\) to which only (successive) movements of the NP or of phrases containing it would be expected to apply.

(11) \([Q_{\text{univ}} \ldots [\text{Dem} \ldots [\text{Num}_{\text{ord}} \ldots [\text{RC} \ldots [\text{Num}_{\text{card}} \ldots [\text{Cl} \ldots [\text{A} \ldots \text{NP]}]})]])

Ideally, all and only the attested orders should follow from the conditions on Merge and the conditions on Move of the type discussed above. Again, more extensive work is needed to check the correctness of this conjecture.

References


\(^{35}\) As shown by the prenominal order Q Dem Num A N of several Indo-European languages (*all those four new jobs*) and the exact mirror-image order N A Num Dem Q, possible in several Semitic languages (Cinque 2000, Shlonsky 2004, Mao Naga (Tibeto-Burman; Giridhar 1994:452), and other languages. There may also be a position of universal quantification below Dem (possibly with distributive meaning). See, for example, the case of Korean discussed by Szabolcsi (1994:sec. 6): *iku motun salan* ‘(lit.) this/the every person’.

\(^{36}\) As in the IP space (see Cinque 1999), certain modifiers may occupy more than one position, giving the impression of a flexible word order across languages. This may be true of possessors, for example (see various contributions in Alexiadou and Wilder 1998), and relative clauses, attracted to complementizers possibly merged (like Ps) in different positions (see Kayne 2000, 2002).


*Dipartimento di Scienze del Linguaggio*  
*Università di Venezia*  
*Dorsoduro 1453*  
*30100 Venice*  
*Italy*  
*cinque@unive.it*