Set-theoretic Constituency & Morphophonological Conditions on Linearization

Andrew Carnie
Acknowledgements

- The work here would be impossible without the help of Muriel Fisher, a native speaker of Scottish Gaelic from Skye who has been working with me for the past year to develop a database of Scottish Gaelic sentence types. Mòran taing, a Mhuirgheal.

- Thanks also to David Adger and Heidi Harley for helpful discussion.

- This work has been supported by a grant from the US National Science Foundation (grant number BCS0639059).
Context

- An interesting consequence of traditional, 2-dimensional trees with constraints against multidomination, line-crossing and multidimensionality is that we are able to construct relatively simple principles of linearization (head parameters, the LCA).

- A challenge to such approaches then are cases where linearization is not transparently based in arboreal terms.
Context

- For example, if we can find examples of linearization phenomena that are subject to lexical, morphological or phonological constraints on their ordering, then that suggests that arboreal linearization principles are overly simplistic.
  - Clitics
  - Heavy NP Shift
  - Focus Inversion

- Problem with all of the above: they all have semantic and syntactic consequences as well as phonological ones
Goal #1

- To demonstrate the existence of an ordering principle that seems to be purely prosodic in nature (although it is sensitive to syntactic structure): Pronoun Post-posing (PPP) in Modern Scottish Gaelic and Modern Irish
Consequence

- It turns out that PPP, applies fairly freely to structures that have adjuncts in them. This will lead me to attempt an adequate set-theoretic description of Adjuncts. Something that has eluded BPS and its near relatives.

- In particular it calls for a looser notion of constituency that allows multi-dimensionality, cast as unrooted (or multiply rooted) graphs found precisely in the case of adjuncts.
Not really a consequence

- The theory of multidimensionality and linearization I propose is:
  - non-arboreal, unrooted or multirooted, multidimensional
  - based on featural dependencies (checking)
  - couched in sets of head-dependent pairs defined along featural grounds
  - constructed to allow for language specific, morpheme specific ordering principles

- This isn’t really a consequence of PPP, nor will I really argue for it, but it’s fun so I’ll do it anyway.
A consequence of the consequence (Goal #2)

- This theoretical approach will in turn allow me to discuss a different morphological condition on linearization, which is the spellout of head-movement chains. I will give a DM style Vocabulary insertion account where elements are linearized in head-movement (actually non-movement AGREE) chains.
Although this talk is couched within a typical generative/minimalist set of assumptions (I have merge, feature checking, phases, etc.), I do stray frequently from official doctrine.

Roman Church vs. High Anglicanism:
- Distributed Morphology
- Conservative functional structure (non-cartographic)

Apostatically & Heretically:
- I use HPSG-like complex-valued features to do feature checking (see Carnie forthcoming).
- Feature checking amounts to relational dependencies (as in dependency grammar, so constituency is an epiphenomenon of linearized dependencies, so the representations are non-arboreal and multidimensional
Outline

- A Phonological Condition on Linearization: Scottish Gaelic PPP
- Dependency based MERGE
- Back to Scottish Gaelic PPP
- The Linearization of Head-Movement Chains
Part 1

- A Phonological Condition on Linearization: Scottish Gaelic PPP
- Dependency based MERGE
- Back to Scottish Gaelic PPP
- The Linearization of Head-Movement Chains
Scottish Gaelic Pronoun Post-Posing (PPP)

(1)

a. Chuir Muirgheal am ball air a’ bhòrd. V S O PP
   Put.past Muriel the ball on the table
   “Muriel put the ball on the table”

b. *Chuir Muirgheal air a’ bhòrd am ball *V S PP O

c. ?Chuir Muirgheal e air a’ bhòrd. ?V S e PP
   Put.past Muriel 3sm on the table
   Muriel put it on the table”

d. Chuir Muirgheal air a’ bhòrd e. V S PP e
Syntactic Accounts

- McCloskey & Chung 1989, rightward movement
- Duffield 1994. 2nd position head: W, hosts clitic pronouns; remnant TP shifts to spec, WP
Problems with syntactic accounts

- Doesn’t allow pronoun groups
- Pronoun can’t appear after embedded clauses (in remnant VP)

1. Chuala mi e air a ràdh [CP gum b’ fhathaist dha a’ bhidh annsin]
   heard I it perf its saying that cop used-to to.him prt be there
   I heard it said that he used to be there” VS e ... CP

2. Chuala mi air a ràdh e [CP gum b’ fhathaist dha a’ bhidh annsin] VS ... e CP

3. *Chuala mi air a ràdh [CP gum b’ fhathaist dha a’ bhidh annsin] e *VS ... CP e
PPP is postsyntactic

- Has no semantic effect
- Adger 2007: VP ellipsis is PF deletion, PPP does not save pronoun
- *Smaoinich mi gun do dh’fhàg mi mo leabhar agus dh’fhàg [\text{VP \text{mi \text{t}}}] aig an sgoil \text{ei}. Thought! that past left I my book and left at the school him ‘I thought I left my book and I had left it at school.’
- *... V [\text{VP-S \text{t}}] PP \text{ei}
Adger’s rule

- Encliticize e-grade pronouns to the non-verb, non-subject word with
  - Focal Stress
  - Nuclear stress

- Adger’s NSR is given on the handout.
Adger’s rule at work

6) Cliticized to nuclear stress

- Chunnaic Mòrag [ann an] Lunnainn [an +dè] i. V S PP Adv i
  ‘Mòrag saw her in London yesterday.’ (Adger 1997)

7) Cliticized to focal stress.

- Q. Càit’ am faic Mòrag do mhàthair?
  where wh-C saw Morag 2s.poss mother
  ‘Where did Mòrag see your mother?’ (Adger 1997)

- A. Chunnaic i [ann an] LUNNainn i [an dè]. V S PP i Adv
  saw 3sf [in] London 3sf [yesterday]
  ‘She saw her in London yesterday.’ (Adger 1997)
Optionality

8)

a) Thug Muirgheal e do Phòl.  
    Give.pst Muriel 3sm to Paul  
    “Muriel give it to Paul”

b) Thug Muirgheal do Phòl e  
    (vaguely preferred)
Avoid iambs

- PPP is obligatory when non-movement would result in an iamb (*dha* is not a pure function word, it is an inflected preposition thus phrasal):

  \[
  \begin{array}{c}
  \times \\
  (x 
  \end{array}
  \]

  a) Thug Muirgheal e dha Dihaoine
  Gave Muriel 3sm to3sm Friday
  “Muriel gave it to him on Friday”

  \[
  \begin{array}{c}
  \times \\
  (x 
  \end{array}
  \]

  b) Thug Muirgheal dha e Dihaoine

  \[
  \begin{array}{c}
  \times \\
  (x 
  \end{array}
  \]
Form a Trochee

10) PPP resists attachment to already existing final trochees (with focal stress on temporal adverb Dihaoine 'Friday').

a) Phòg Muirgheal Pòl Dihaoine
   Kissed Muriel Paul Friday
   “Muriel Kissed Paul on Friday”

b) Phòg Muirgheal e Dihaoine V S e Adv

c) *Phòg Muirgheal Dihaoine e *V S Adv e

d) ?Phòg Muirgheal Dihaoin’ e ?V S Adv e
Form Trochee

11) More examples of resistance to already existing trochees with clear VP adverbs

a) Leugh Bob an leabhar gu cùramach/gu tric
   Read.past Bob the book adv careful/ adv frequent
   “Bob read the book carefully/frequently”

b) Leugh Bob e gu cùramach
   V S e Adv

c) ?Leugh Bob gu cùramach e
   ?V S Adv e

d) ?Leugh Bob e gu tric
   ?V S e Adv

e) Leugh Bob gu tric e
   V S Adv e

warning: judgments are murky and nebulous
12) More examples of resistance to already existing trochees, this time with a choice of a monosyllable or a bisyllabic trochee to attach to.

a. ?Chic Muirgheal e gu cùramach gu Pòl e V S e Adv PP
   kick.past Muriel 3sm adv careful to paul
   “Muriel kicked it to Paul carefully”

b. Chic Muirgheal gu cùramach gu Pòl e V S Adv PP e (preferred)

c. (?)Chic Muirgheal gu cùramach e gu Pòl (?)/V S Adv e PP (dispreferred)
12) same facts but with order of the PP and Adv reversed

a) ?Chic Muirgheal e gu Pòl gu cùramach ?V S e PP Adv
Kick.past Muriel 3sm to Paul 3sm adv careful
“Muriel kicked it to Paul carefully” (ok, but not preferred)

b) Chic Muirgheal gu Pòl e gu cùramach V S PP e Adv (preferred)

(c) ?Chic Muirgheal gu Pòl gu cùramach e ?V S PP Adv e
(ok, but requires focal stress on cùramach)
13) Vowel Vowel sequences seem to trigger PPP

a. ?Bhris e e leis.  
   Break.past 3sm 3sm with.3sm  
   "he broke it with it"

b. Bhris e leis e  
   V S PP e

warning: judgments are murky and nebulous
13) Vowel Vowel sequences seem to trigger PPP

C. ?Chuir Muirgheal e air a’ bhòrd ?V S e PP
   Put Muriel 3sm on the table
   “Muriel put the ball on the table

D. Chuir Muirgheal air a’ bhòrd e V S PP e
13) PPP seems to avoid creating VV sequences (e-g)

e. Dh’fhàgar e ‘na laighe air an lár  
past’give.impers 3sm in.3sm.poss lie.vn on the ground  
“It was left lying on the ground.

f. *Dh’fhàgar ‘na laighe e air an lár  

g. Dh’fhàgar ‘na laighe air an lár e
iad vs ‘ad

- There are three pronouns that shift: e, i, iad. The last of these is trochaic already: /'i.ət/. When it shifts to adjoin to the stress it must take its reduced form 'ad /ət/.

  a) Dh’inns i iad do Sheumais past’say 3sf 3pl to James “She said them to James.”

  b) Dh’inns i do Sheumais ‘ad

- This shows us two things.
  - The motivation for the movement has something to do with the pronoun itself (only the 'ad form postposes)
  - The motivation is clearly prosodic rather than syntactic.
Projection of Verb

- Phonological conditions aside, pronoun post-posing seems not to care which projection of the V it cliticizes to (Data from Modern Irish):
  - Deir sí gur bhris sé an doras leis an ord aréir
    - says she that broke he the door with the hammer yesterday
    - She says that he broke the door with the hammer yesterday.
  - ?Deir sí gur bhris sé é leis an ord aréir
  - Deir sí gur bhris sé leis an ord é aréir
  - Deir sí gur bhris sé leis an ord aréir é
Descriptive Generalization

- Optionally linearize object pronouns to a modifier of the verb bearing Focal stress or Nuclear stress with a preference to forming trochees, and avoiding VV sequences.

- Syntactic conditions: May not adjoin after a tensed embedded clause. Must adjoin to a modifier of the verb (or temporal adverb when focused).
An aside

- There is a lot of the data on SG PPP that I still don't understand, including some very subtle judgments and contradictory data. You can see transcripts of my sessions with my native speaker on http://www.dingo.sbs.arizona.edu/~gaelic, click on "data" and then on "Andrew Carnie: Basic Clause Structure". Comments and reanalysis are welcome.
Consequences

- Linearization is post-syntactic and thus open to phonological conditioning.
- Linearization can't be (purely) based on asymmetric c-command.
- The phenomenon here seems to treat all adjuncts as equals for linearization of object pronouns, once we factor out the phonological conditions. How is this possible?
A bulldozer to plant a flower?

- This is where a looser theory of constituency comes in. I’ll sketch such a theory next—it’s more complicated than is necessary for dealing with the SG facts, but I wanted to set out a theory that is relatively complete for concreteness.
Part 2

- A Phonological Condition on Linearization: Scottish Gaelic PPP
- Dependency based MERGE
- Back to Scottish Gaelic PPP
- The Linearization of Head-Movement Chains
Intellectual Antecedents

- van Reimsdijk, Haegeman, Pesetsky, etc.
- Brody 2000, Bury 2005
- Tesnèire 1959, Mel’čuk 1988, Hudson 1994
- and others
Feature driven Merge

Zwart and others

15) Merge: given a Numeration \{\alpha, \beta\}, where \alpha has an unchecked SUBCAT feature and \beta checks that feature, the resultant merge is the ordered pair of terminals \langle \alpha, \beta \rangle \ (read as "\beta is the dependent of \alpha")

Note:
- No labels
- No domination (the only relations exist between terminals).
Example 1

resultant merge: <THESE, CATS>
Example 2

KISS

\[
\begin{bmatrix}
\text{INTERNAL} & D_{(+\text{animate})} \\
\text{EXTERNAL} & D
\end{bmatrix}
\]

- Assume the sets <THE, CAT> and <A, POODLE> have already been merged.
- Assume that the "the" in <THE, CAT> bears an animacy feature.

<KISS, THE> this satisfies the EXTERNAL feature
<KISS, A> this satisfies the verb's INTERNAL feature.
sets of pairs are relations

17) Following Collins and Ura (2004), let us assume that dependencies like those in (15) and (16) express relations. Unlike Collins and Ura let us define those relations in terms of the feature values that are being checked.

- S1 = {R1, R2}
  - R1: INTERNAL = { <THE, CAT>, <A, POODLE>, <KISS, A> }
  - R2: EXTERNAL = { <KISS, THE> }
18) First pass at a linearization rule for English:
   a) If \( <x, y> \in \text{INTERNAL} \), then \( x \prec y \)
   b) If \( <x, y> \in \text{EXTERNAL} \), then \( y \prec x \)

19) Spellout:
   a. Let \( Z \) be a phase, let \( X \) be a head, and let \( Y \) be a dependent with no unchecked features, linearize using language specific linearization rules, resulting in either \( /X+Y/ \) or \( /Y+X/ \) in \( Z \).
   b. Once a precedence or adjacency relation is established at Spell-Out, it is not disrupted later in the derivation (Collins & Ura 2004’s version of non-tangling)
The cat kissed a puppy

i. THE ≺ CAT   (18a)
ii. /the+cat/   (19a)   DP Phase 1
iii. A ≺ POODLE (18a)
iv. /a+poodle/  (19a)   DP Phase 2
v. KISS ≺ A    (18a)
vi. KISS ≺ /a+poodle/ by (19b) and (v)
vii. /kissed+a+poodle/ by (19a)   VP Phase
viii. THE ≺ KISS (18b)
ix. /the+cat/ ≺ KISS by (19b) & (viii)
x. /the+cat/ ≺ /kissed+a+poodle/ (19b), (viii), (ii), and (ix)

- (tense isn’t specified here so the linearization is not complete)
With “will”

- **WILL: SUBCAT** [+Aux, EXTERNAL [Case nom], INTERNAL V]
- Let us assume that the THE in “the cat” is marked as nominative case
- **S2 = <R1, R2>**
  - R1: INTERNAL = { <THE, CAT>, <A, POODLE>, <KISS, A>, <WILL, KISS> }
  - R2: EXTERNAL = { <KISS, THE>, <WILL, THE> }
With “will”

- Linearization of S2:
  
  x. /the+cat/ ≺ /kiss+a+poodle/ from previous linearization
  xi. WILL ≺ KISS (18a)
  xii. WILL ≺ /kiss+a+poodle/ (19b), (xi)
  xiii. THE ≺ WILL (18b)
  xiv. /the+cat/ ≺ WILL (19b), (xiii)
  xv. /the+cat+will+kiss+a+poodle/ (19a), (x-xiv)
Adjuncts: three properties

A. Adjuncts have scope with respect to one another

i) The sexy muscled firemen
ii) The muscled sexy firemen
(The subset of muscled firemen who are sexy vs. The subset of sexy firemen who are muscled)

iii) The boxes in the storeroom with the red label
iv) The boxes with the red label in the store room
(The subset of the boxes in the storeroom with red labels, not the ones with the yellow labels vs. the subset of boxes with red labels that are in the store room, not the ones with red labels in the office.)
Adjuncts: three properties

B. Adjuncts are not selected for by the head of the phrase they modify, but instead select for that head.

C. Adjuncts don’t participate in the main spine of tree with respect to c-command and reconstruction.
Claim: adjuncts have precisely the property that they select for dependencies to scope over.

We will encode this in a MOD feature. Where Ⓥ and Ⓦ stand for “projection of v” and “projection of n” respectively. These are projections of dependencies.

- skillfully [Mod Ⓥ]
- yellow [Mod Ⓦ]
- in [Mod Ⓦ/_codec]
Projection

27) Projection: (A recasting of Speas 1990 in terms of dependencies)

- An ordered pair ⊗ is a projection of x if ⊗ contains an uninterrupted sequence of left angle brackets between the brackets defining ⊗ and x.
Projection

- ① is a projection of a in the following:
  - ① = <a, b>
  - ① = <<a, b>, c>
  - ① = <<<a, b>, c>, d>

- But not in the following cases:
  - ① ≠ <b, a>
  - ① ≠ <<b, a>, c>
  - ① ≠ <b, <a, c>>
  - ① ≠ <<b, <a, c>>, d>
29) The rule adjoin:

- ADJOIN only applies in the context of a mod feature. Given a lexical item y whose MOD feature selects for a dependency-projection $\times$, create an ordered set headed by $\times$: $<\times, y>$
Linearization in MOD

- if $<x, y> \in \text{MOD}$, and $y = \text{GENERAL}$ then $x < y$ (English)
- if $<x, y> \in \text{MOD}$, and $y = \text{SEAN}$ then $y < x$ (SG)
- if $<x, y> \in \text{MOD}$, and $y = \text{Adj}$ then $y < x$ (English)
- if $<x, y> \in \text{MOD}$, and $y = \text{Adj}$ then $x < y$ (SG)
- if $<x, y> \in \text{MOD}$, and $y = P$ then $x < y$ (English and SG)
The cat kissed a puppy in the woodshed

- $S_3 = \{<R1, R2>, R3\}$
- $R1: \text{INTERNAL} = \{<\text{THE}_1, \text{CAT}>, <\text{A}, \text{PUPPY}>, <\text{THE}_2, \text{WOODSHED}>, <\text{KISS}, \text{A}>, <\text{IN}, \text{THE}_2>, <\text{WILL}, \text{KISS}>\}$
- $R2: \text{EXTERNAL} = \{<\text{WILL}, \text{THE}_1>, <\text{KISS}, \text{THE}_1>\}$
- $R3: \text{MOD} = \{<\text{KISS}, \text{A}>, \text{IN}>\}$
A more complicated example using two adjuncts can be found on your handout in (31e)
Linearization of S3

- **DP Phase:**
  - THE1 < CAT
  - /the+cat/

- **DP Phase:**
  - A < PUPPY
  - /a+puppy/

- **PP/DP Phase:**
  - THE2 < WOODSHED
  - /the+woodshed/
  - IN < THE
  - IN < /the+woodshed/
  - /in+the+woodshed/

- **VP phase**
  - KISS < A
  - KISS < /a+puppy/
  - /kiss+a+puppy/
  - <KISS, A> < IN
  - /kiss+a+puppy/</in+the+woodshed/
  - /kiss+a+puppy+in+the+woodshed/

- **TP phase (includes VP edge)**
  - THE < KISS
  - /the+cat/ < KISS
  - /the+cat/ < /kiss+a+puppy+in+the+woodshed/
  - THE < WILL
  - /the+cat/ < WILL
  - WILL < KISS
  - WILL < /kiss+a+puppy+in+the+woodshed/
  - /the+cat+will+kiss+a+puppy+in+the+woodshed/
Part 3

- A Phonological Condition on Linearization: Scottish Gaelic PPP
- Dependency based MERGE
- Back to Scottish Gaelic PPP
- The Linearization of Head-Movement Chains
PPP

- Chuir Muirgheal t\textsubscript{put} air a’ bhòrd e
  - Put Muriel on the table 3sm

- R1: INTERNAL = \{ ... , \langle t_{\text{put}}, E \rangle, \langle \text{AIR}, A' \rangle, \langle A', \text{BHÒRD} \rangle \}

- R3: MOD = \{ \langle \langle t_{\text{put}}, E \rangle, \text{AIR} \rangle, \}
The VIs for accusative e, i, and ‘ad (but not iad or any other noun) have the following:

\[
x \rightarrow (x \ x)
\]

\[
<V, E> : <\text{\textcircled{\text{v}}}, /\ldots \sigma\text{c#} +\ldots>/
\]

(where \(\sigma\text{c#}\) stands for a closed syllable at the end of the dependency)

Additional stipulations are needed to make sure that the stresses in embedded clauses aren’t accessible to the rule.

Applies to e, i, ‘ad in SG, and to all 3rd person pronouns in Irish (including prepositionally marked ones): Clearly a lexical stipulation.
air a’ bhòrd e

i. V < E

ii. */Ø+e/  <-- not an ideal linearization because of clitic.

iii. <<V, E> AIR> = \( \text{V} \)

iv. 

\[
\begin{array}{c}
\text{x} \\
\text{(x}
\end{array}
\]

<Ø, E> < /air+a+bhòrd/

v. 

\[
\begin{array}{c}
\text{x} \\
\text{(x x)}
\end{array}
\]

Ø /air+a+bhòrd+e/

- And similarly to any other adjunct that is the dependent in a projection to the verb.
...iad air a bhòrd

- V < IAD
- /Ø+iad/
- <V, IAD> < AIR
- /Ø+iad/ < /air+a+bhòrd/
- /Ø+iad+air+a+bhòrd/
Part 4

- A Phonological Condition on Linearization: Scottish Gaelic PPP
- Dependency based MERGE
- Back to Scottish Gaelic PPP
- The Linearization of Head-Movement Chains
Create a dependency between two terminals \( <x,y> \), such that \( y \) is an uninterpretable checker for the first’s Head features.

C-command condition: \( \exists R_n, R_n \subseteq \text{INTERNAL} : R_n = \text{a sequence of connected dependencies}, \text{where} \ x \text{ is the highest head and} \ y \text{ is the lowest dependent.} \)

- LEAVE \ [\text{head} \ uT] 
- PAST \ [\text{head} \ T] 
- R4: Head = \{ <\text{PAST}_T, \text{LEAVE}_{uT} > \}
Connected Dependencies

- Connected dependencies are ones where the dependent of one pair is the head of another.
- A sequence of connected dependencies is set of connected dependency pairs, where each terminal occurs exactly twice (no more and no less), once as the head of a pair and once as the dependent.
Spellout of HEAD dependencies

- Head dependencies are not linearized (linearization happens only in the Subcat relations). But impose the following conditions on spellout:

  - Condition on the spellout of HEAD dependencies: Only 1 element in any sequence of connected head dependencies may be realized by vocabulary insertion. All other elements in the sequence are rendered as null. (similar to Brody 2000)

  - Principle of Ø insertion for uninterpretable features. If at the end of a phase, no vocabulary item exists compatible with an the features on a terminal, then Ø is inserted.
Language-specific VIs

- The choice of where in the head chain a word is inserted is determined by language-specific morphological stipulations on vocabulary. (Similar to Kroeger 1997, Brody 2000)
  - *eats* is of category V (with T features specified)
  - *eat/eaten* are of category V
  - *mange* is of category T
  - *mangé* is of category V.
- In French T is linearized by the lexical items like *mange*, and V is rendered as Ø. This is licensed by the head dependency.
- In English the T terminal is rendered as Ø, and *eats* is inserted into V.
Language-specific VIs

- Prediction: particular lexical items can vary from the normal patterns: Scottish Gaelic Future and past tense verb forms are of category T, the verb be, is of category T in the present tense. There are no other present tense verbs. The present tense must be indicated periphrastically -- where a participial form (with no T features) is inserted into V, and T is realized by an Auxiliary.

- By contrast, Irish allows present tense forms of all verbs. So the present tense involves insertion into T in all forms.
HMC/Intervention effects

- Intervention/Relativized Minimality/HMC effects are due to linearization paradoxes. (cf. Bury)
  - Internal = { … < T, Neg>, <Neg, V> …}
  - Therefore \( T < Neg < V \)
  - But at linearization the \( V \) is in \( T \), so in effect you also get \( V < Neg \).
- Since you have both \( Neg < V \) and \( V < Neg \), you have an unlinearizable string.
Scottish Gaelic exhibits a phonological condition on linearization.

This condition requires a more flexible notion of constituency.

Perhaps with a bit of overkill, a version of generative dependency is given, which allows for language-specific and lexical-item specific principles of linearization, such as SG PPP.

This in turn opens up the possibility of explaining head-movement in terms of dependencies paired with language specific morphological conditions on position of exponence in the syntax.