Today

§1 An overview of the origins and development of head movement.

§2 Chomsky, and head movement a ‘phonological’ operation

§3 Some more alleged empirical properties of head movement.

§4 An aside on some phenomena we won’t be talking about.

1 Origins

Head movement has been around in some form since the mid to late 1970s. Generally speaking, it has been used to derive the positions of verbs cross-linguistically.

1.1 Emonds

- The earliest citations for something like head movement are usually to work by Joe Emonds; Emonds 1978 is fairly representative, proposing V<sup>0</sup>-to-T<sup>0</sup> movement.

- Building on work by Kayne (1975), Emonds is particularly interested in understanding the status of French auxiliaries and the structures in which they occur.

- In particular, Kayne observes various adverbs must precede infinitives, but follow tensed verbs. Emonds notes this is true of negation, as well.

  (1) Jean (n’) aime pas Marie.
  John ne likes NEG Mary

  (2) Ne {pas} posséder {*pas} de voiture en banlieue rend la vie difficile.
  ne NEG own.inf of car in suburbs makes the life difficult

- Emonds proposes that finite verbs undergo movement to tense (X being an adverb or negation in the transformation below):

  (3) Finite Verb Raising:
  Conditions: X doesn’t contain NP or V; obligatory application

- Emonds proposes this rule in the context of a particular view of auxiliary verb structure, which he argues to be left-branching (exploiting the then-burgeoning X<sup>′</sup>-theory). It is actually a part of an argument about the correct placement of tense at deep structure.

- Like many transformations of the time, this rule affects a linear order. The leftmost verb (term 4 above) will be the one to undergo movement to tense (term 2).
This is not really head movement in the sense that we recognize it today. The terms of the transformation don’t distinguish between heads and phrases, we can move anything anywhere as long as it is a constituent, and it relies on linear order to get the correct verb to move.

The novel theoretical insight of this rule is that it resolves an apparent problem of word order by proposing that tense is external to the VP and that verbs move to tense in French.

1.2 Travis

An important development in head movement came in Lisa Travis’s (1984) dissertation. In Chapter 3, Travis’s main goal is understanding various differences between the word orders of German, Dutch, and Yiddish. Much of this discussion centers on the idea that IPs in German and Dutch are left-headed.

The problem she faces is that verbs in German seem to appear in three different places: In Comp, in Infl (between subjects and objects), and in VP.

A key part of understanding this comes in the form of the Head Movement Constraint (HMC):

(5) Head Movement Constraint:
    An X° may only move into the Y° which properly governs it.

This has the effect of linking head positions together. Verbs move to I°, and then to Comp – there is a feeding relationship between these positions.

Thus in a series of verbs, only the highest verb will move to Infl (6). The inflected verb (the one that moves to I°) will be the only one available to move to Comp (7).
The hmc is a notable move toward generalization: We are no longer just talking about verb movement, but movement of syntactic elements in particular syntactic positions.

The hmc remains a central empirical observation to this day. Roberts (2011) recasts this in more modern terms:

(8) **Head Movement Constraint (hmc):**
Head movement of X to Y cannot skip an “intervening” head Z.

(Z intervenes between Y and X iff Y asymmetrically c-commands both X and Z, while Z asymmetrically c-commands X)

Exactly why the hmc should hold, however, is an open and problematic question (as we will see).

Rizzi (1990) has argued that the hmc is a special case of Relativized Minimality, suggesting that it may fall out from more general syntactic locality principles.

### 1.3 Head movement in GB

- In Government and Binding theory, head movement can be summarized as:

(9) Head movement is the case of Move-\(a\) where \(a\) is \(X^\circ\).

- Head movement, being a reflex of Merge-\(a\), was therefore a core syntactic operation.

- As such, it was thought to be constrained by general constraints on movement:

  a. **Structure preservation:**
  In traditional \(X\)-theory, this meant that only heads could move to head positions and only phrases could move to specifiers. This basically required head-to-head adjunction.

  b. **Locality:**
  The locality condition on head movement was the **hmc** (8). The hmc, as mentioned above, forces head movement to be cyclic, since a moving head must pass through each head position on the way to its target.
c. Well-formedness of conditions on traces:
This, of course, means the Empty Category Principle (ecp). This meant that head movement had to proceed out of complements, and that ‘downward’ head movement was not allowed.

• The GB view of head movement survived into early versions of Minimalism, and is probably the formulation that most non-specialists are familiar with.

2 Chomsky on head movement

• As with so many other things, the view of head movement began to shift after the adoption of the Minimalist Program.

• Chomsky (2001:37) suggests ‘the possibility that V-raising is [...] not part of the narrow-syntactic computation but an operation of the phonological component’. He goes on to claim that ‘[t]here are some reasons to suspect that a substantial core of head raising processes, excluding incorporation in the sense of Baker (1988), may fall within the phonological component’. He enumerates the following reasons:

(i) Interpretation of raised verbs:
Verbs that undergo movement to T° or to C° receive the same (LF) interpretation as those verbs which remain in situ.

(ii) The motivation for movement:
It is frequently assumed that a strong [v] feature on T° drives V°-to-T° movement, but this cannot be satisfied by VP movement to SpecTP. The theoretical apparatus driving movement, however, cannot distinguish between head and phrasal movement, so both should be possible. The ‘phonetically affixal character of the inflectional categories’ suggests head movement is a PF process.

(iii) Differences between head movement and phrasal movement:
   a. It is an adjunction rule.
   b. It is countercyclic.
   c. A moved head does not c-command its trace.
   d. It is problematic in ways similar to feature movement.
   e. It observes different locality conditions.
   f. It is not successive cyclic.

2.1 The empirical claim: Interpretive effects

• Notice that only point (i) is a properly empirical claim; the rest are theory-internal concerns. As a consequence, probably, point (i) is the most contentious claim in modern theorizing on head-movement.

• If head movement feeds an interpretive difference, then the simplest conclusion is that head movement is a narrow syntactic operation.

• If head movement can be shown to have interpretive effects, we can be relatively confident that head movement is part of the narrow syntax. Then the theory of movement would need to be adjusted to accommodate the issues raised in points (ii) and (iii).
• Note, however, that a lack of interpretive effects is not evidence that head movement is a PF operation; it is merely evidence consistent with that view.

• Syntactic movements often lack interpretive effects for a host of reasons, and there are arguably reasons that head movement might lack them. This is why the theory-internal concerns in (ii) and (iii) are of importance.

2.2 Motivation: Feature driven movement

• Agree is a central theoretical mechanism in Chomsky 2001, so making it function consistently is a key motivation for Chomsky here.

• The typical Minimalist assumption, given Bare Phrase Structure, is that the a head projects the same label as the phrase it forms. The feature make-up of a phrase is the same as the head of the phrase – there is no way to distinguish between a head and the phrase that projects from it.

• Consequently, there is no way for Agree distinguish between phrases and heads using features.

• If all syntactic movement is feature-driven, and \( V_0 \)-to-\( T^o \) movement is driven by a strong \([v]\) feature on \( T^o \) (Pollock 1989), VP movement as in (11) should be just as acceptable as \( V_0 \)-to-\( T^o \) movement, as in (10), since both should check the \([v]\) feature.

\[
(10) \quad \begin{array}{c}
TP \\
\downarrow \\
T^o \quad vP \\
\downarrow \\
V^o \quad [v] \quad t_v \\
\end{array}
\]

\[
(11) \quad \begin{array}{c}
TP \\
\downarrow \\
VP \quad T' \\
\Downarrow \\
\vdots \quad T^o \quad vP \\
\end{array}
\]

• As Preminger (2016) points out, this would seem to suggest that head movement should not even be possible. Locality conditions on Agree mean that VP is a nearer goal for the probe on \( T^o \) than \( V^o \), so VP movement should be preferred to V movement. Only (11) should be possible.

• If head movement is not (feature-driven) movement, then these derivations are no longer in competition.

2.3 The theory of movement: The extension condition

• The list of issues in (iii) are several issues that violate contemporary views of movement in Minimalism.

• Several of the points in (iii) are related to the Extension Condition of Chomsky (1995), namely:

  a. It is an adjunction rule.
  b. It is countercyclic.
  c. A moved head does not c-command its trace.

• The Extension Condition requires that Merge (and therefore, Move) apply at the topmost node of the tree:
(12) **Extension Condition:**
   Merge should be effected at the root.

   – In order for a head $X^o$ to adjoin to another head $Y^o$, $Y^o$ must already have merged, and so $X^o$ cannot adjoin at the root.
   – The Extension Condition is ruled out countercyclic movement, which head-to-head adjunction would be a case of.
   – The Extension Condition also derives the fact that a moved phrase c-commands its trace, since it will always be a sister to the constituent out of which it moves.

• The point about the locality conditions (iii-e) is more equivocal, since the HMC can be explained as a special case of Relativized Minimality.

• The fact that it is not successive cyclic is related to the fact that it is impossible to extract from a complex head (as I discuss below).

2.4 Discussion

The real issue with points (ii) and (iii) is that they are completely theory-internal.

• If we construct Agree differently, or if we reconceptualize how features are projected from heads to phrases, point (ii) might be avoided. If we reject the Extension Condition, then many of the issues in point (iii) might dissolve.

• Question: These theoretical mechanisms were not necessarily made with head movement in mind. Should they have been?

   – Imagine if Chomsky came to this problem with the intent of keeping head movement as part of the narrow syntax. Could there have been a way to build the theory to meet that end?
   – Here we have a theory that was designed to handle phrasal movement. Head movement is thought to behave differently from phrasal movement. Is the fact, then, that it does not fit well with a theory meant to describe phrasal movement sufficient grounds for excluding it from the narrow syntax?
   – This is why point (i) is such a contentious issue – the empirical evidence supporting this move is actually fairly thin.

• There’s also a big gap left in the theory if head movement is not syntactic.

   – Heads are syntactic objects. How do we account for the fact that head movement appears to target the head position in a syntactic structure? What role should the syntax play?
   – The HMC looks a lot like a syntactic constraint (especially given its outward resemblance to Relativized Minimality). Any non-syntactic account of head movement must reckon with the fact that head movement takes place between adjacent heads and that any intervening non-head material (adjuncts, specifiers) can be skipped.
– If there is ‘phonological’ head movement, we should wonder whether and why
head movement is not available in the syntax. Merge must to be able to take
bare heads as its terms (since the contents of a numeration is a collection of
lexical items). Internal Merge should also be able to do so in principle.

• I raise these questions because I think a healthy dose of skepticism is warranted here.
The empirical evidence here is not particularly strong; in general, more evidence
should be brought to bear on these issues.

3 Additional properties of head movement

• There are several other ways in which head movement is reported to differ from
phrasal movement.

• The following list, which overlaps with Chomsky’s (2001) in several places, is from
Platzack (2013), who identifies several additional properties we have not seen yet.

a. Violation of the extension condition
HM does not extend the root when moved, contrary to phrasal movement.

b. Locality
HM is restricted to occur within a single extended projection [of the verb –NL].
Phrasal movement is not restricted in such a way.

c. Relativized minimality
HM but not phrasal movement is blocked by an intervening head.

d. No extraction
If a head $\beta$ moves to $\alpha$, then $\{\alpha + \beta\}$ acts as one constituent. A moved phrase may
display a similar “freezing” effect.

e. The higher the bigger
If $\{\alpha + \beta\}$ is the result of HM of $\beta$ to $\alpha$, then the features of $\beta$ are a proper subset of
those of $\alpha$. There is no similar effect of phrasal movement.

f. Mirror Principle
HM obeys the Mirror principle, saying that the order of affixes attached to a head is
a mirror image of the order of functional heads corresponding to these affixes. There
is no mirror principle in connection with phrasal movement.

g. No semantic effects
HM seems to lack semantic effects, whereas phrasal movement may have such ef-
effects.

• Of these, only b, d, e, and f are completely novel, so I will discuss them below.

3.1 Locality (b)

• Platzack claims that ‘a head cannot head move out of its maximal functional pro-
jection, hence V cannot head move out of CP’, contrasting verb movement with
topicalization:
(13) This book, John thinks that Eva read last summer.
(14) a. *Has Mary said that John stolen the book?
b. Has Mary said that John stole the book?

• I’m not totally sure what this is meant to show.
  1. If this is meant to show that head movement head movement cannot be trig-
     gered long-distance, than this is already covered by the HMC.
  2. If this is meant to show that heads cannot undergo topicalization (in English),
     then this is not really head movement in the strict sense.

• The way he puts it, though, would seem to suggest that head movement is clause-
  bounded. If this is right, then there shouldn’t be anything like like C°-to-V° move-
  ment. I’ve never seen this claimed anywhere, and also I don’t know of any clear evi-
  dence that this does happen.

3.2 No extraction (d)

• Sometimes referred to as the ban on EXCORPORATION, it is thought that it is impos-
  sible to move out of a complex head.

• The result, then, is that if one wants to move a head X° to Z° across an intervening
  head Y°, then one must also move Y° to Z°.
    – The HMC precludes straight-out skipping over Y°, so X° must move to Y° first.
    – It seems, though, that subsequent movement of X° requires movement of Y°
      as well. It is not possible to move X° without Y°.

• Roberts (2003, 2011) points out that nothing in the GB conception of head move-
  ment explains the ban on excorporation; Platzack (2013) notes that it remains a stip-
  ulation.

3.3 The higher, the bigger (e)

• A consequence of the inability to extract from a complex head is that the iterated ap-
  plication of head movement results in progressively larger syntactic elements. Phrasal
  movement does not really do this.

• A related observation, though, is that this generalization is not necessarily diagnostic
  of head movement having occurred, due to confounds like AFFIX HOPPING.
    – Verbs are not typically reckoned to move past v° in English. Despite this, they
      appear with tense and aspect affixes associated with a higher head.
    – Moreover, T°-to-C° movement in most cases does not result in increased mor-
      phological complexity. Thus, ’bigger’ can, at best, only mean syntactically bigger.
3.4 The mirror principle (f)

- The Mirror Principle of Baker (1985)

\[(15)\] *The Mirror Principle:* Morphological derivations must directly reflect syntactic derivations (and *vice versa*).

- In morpho-syntactic theories where word-building is the result of syntactic operations (read: head movement), the morphological structure of the word will reflect the syntactic structure.

- For instance, if a verb acquires affixes by moving through various functional heads, the order of the heads will reflect the order of the functional heads through which it moved.

\[(16)\]
```
TP
  \[\text{v}^0\] 
  \[T^0\]
  \[v'\]
  \[\sqrt{\text{sing}}\]
  \[-a\]
  \[-n\]
```

- This should fall out from the HMC and the ban on excorporation.

4 Important things we are not talking about extensively

4.1 Incorporation

- **Incorporation** (Baker 1988) is a rich and complicated topic. Very broadly speaking, incorporation occurs when the complement of a verb appears as a proper morphological subpart of that verb.

\[(17)\] Incorporation in Onondaga (Iroquoian).

a. Pet \[wâ?-\text{htu}-\text{?t}-a\] ne\[e\] o-hwist-.
   Pat \[PAST\text{-3MS-lost-CAUS-ASP}\] the \[\text{PRE-money-SUF}\]
   ‘Pat lost money.’

b. Pet \[wâ?-\text{hwist-htu}-\text{?t}-a\]?
   Pat \[PAST\text{-3MS-money-lost-CAUS-ASP}\]
   ‘Pat lost money.’

- The traditional view of this is that this is derived via head movement.
Head movement: An overview

Historically, incorporation has played a central role in the theory of head movement. Reading Travis (1984), for instance, one gets the sense that there was no meaningful distinction between the two.

However, it has taken a backseat in recent discussions on the nature of head movement, quite possibly because Chomsky (2001) himself brackets it.

This is certainly a mistake, and one that I am unfortunately propagating here.

4.2 Clitics

Clitics, in some cases, look a lot like individual heads

For example, Uriagereka (1995) argues that Galician object clitics are heads that adjoin to a functional position to which the verb also moves.

These clitics originate in argument positions. While the verb movement to F0 follows the HMC, movement of the clitic to F0 does not.

Clitics are fascinating, but well beyond the purview of what we are talking about here.
References


den Besten, Hans. 1977. On the Interaction of Root Transformations and Lexical Dele-
tive Rules. Ms.


Preminger, Omer. 2016. What the PCC tells us about "abstract" agreement, head movement, and locality. URL http://ling.auf.net/lingbuzz/003221, manuscript, University of Maryland, December 2016.


