1. HEAD-MOVEMENT: THE EMPIRICAL PHENOMENON

Fact: two syntactically distinct terminals, sometimes linearly separated by overt phonological material, may end up in the position of the higher of them as one unit for the purposes of both syntactic and morphological operations:

(1) a. XP
   Y°
   \* Spec
   Y
   \* WP

   b. XP
   X°
   \* Spec
   Y°
   X
   \* t
   \* WP

Core phenomena: French finite verbs like fin-
ir-a ‘finish-FUT-3SG’ (V-T-Agr) or (although some would disagree) English synthetic comparatives like tall-er (A-Deg).

2. HEAD-MOVEMENT AS A PF PHENOMENON

Chomsky 2000: head-movement (with the possible exception of incorporation) takes place at the PF branch of the derivation, after Spell-Out.

NB Incorporation has somehow fallen between the cracks, but one thing is clear: there are several kinds of it, some of which (e.g., noun incorporation in West Greenlandic, see Rosen 1989) occur in configurations that do not allow head-movement.

The exclusion of movement of heads from narrow syntax is unmotivated and it can be shown that head-movement cannot be triggered by purely phonological factors (Matushansky 2006).

Prediction: head-movement has no semantic effects

2.1. LF effects of head-movement

Matushansky 2006: most items moved by head-movement are predicates of some sort (verbs, nouns, most affixes, etc.), and therefore obligatorily reconstruct:

(3) t = \[ \[ \text{dances} \] \] (Bella)
   ↓
   \[ \[ \text{dances} \] \] (Bella)
   \[ \[ \text{dances} \] \] (Bella)
   = \[ \[ \text{dances} \] \] (Bella)
   = \[ \[ \text{dances} \] \] (Bella)
   = \[ \[ \text{dances} \] \] (Bella)

Only head-movement of a quantificational head can give rise to LF effects:

(4) a. Yolanda can’t leave.  ¬
   b. Yolanda shouldn’t leave.  ¬

(5) a. Passengers may not speak to the driver.
   b. You may not think so at first, but it is a very smart rule.

Important: intensional verbs that don’t raise do not outscope negation.

Lechner 2005: scope-splitting shows that head-movement has semantic effects:

(6) a. Not everyone can be an orphan.  ¬\[ \forall \]
   b. Sam can find no solution.  ¬\[ \exists \]


(7) a. Every critic seemed to like the movie.
   b. It seemed that every critic liked the movie.

Negative indefinites also don't reconstruct (Lasnik 1999, Wurmbrand and Bobaljik 1999, von Fintel and Iatridou 2004):
(8)  a. No critic is certain to like the movie.
    b. It is certain that no critic will like the movie.

The narrow-scope readings of (6) have to result from the modal taking wide scope (by head-movement).

NB Lechner 2005 assumes that negative universals and negative indefinites involve a wide-scope abstract NOT operator licensing overt negation in what really are universals and indefinites respectively (von Stechow 1993, Penka 2002), but is it necessary to appeal to negative universals? Isn't the same point made by the interpretation of Everyone can't be happy or Every student mustn't get an A (Sauerland 2003)? Why don't you get the narrow scope of the universal in Everyone can survive on the food we have? Observe that Lechner's explanation contradicts that in latitudio and Zeijlstra 2010: modals reconstruct unless they can't.

Roberts 2010 (building on McCloskey 1996): NPIs can be licensed by head-movement:

(9)  a. *Anyone didn't see John.
    b. John didn't see anyone.

(10) a. *Which one of them does anybody like?
    b. Which one of them doesn't anybody like?

Overt head-movement of negation to a position c-commanding the NPI in the subject position licenses it.

This, however, is fallacious reasoning, as "Progovac 1994: 97-98 observes that polarity items can be licensed in wh-questions, but these must be rhetorical questions" (Roberts 2010)

(11) When did Mary insult anyone?

Negative questions are usually rhetorical.

(12) a. I know which book everyone read. *Which book didn't anyone read?
    b. Couldn't she look for someone tomorrow?

When negation is interpreted as negation, an NPI in the subject position is not licensed.

This is not surprising: the semantic type of negation ensures semantic reconstruction. Summary: only modals provide any evidence for LF effects of head-movement.

2.2. The mechanics of PF head-movement

Consider verb-raising in a structure like (13):

(13) When does V^0→v^0 occur? Unless PF can handle two phases at once or the notion of the phase edge is reintroduced at PF, PF movement across a phase boundary is impossible.

2.3. Some specific proposals

Hale and Keyser 2002, Harley 2005: Conflation is an operation occurring at Merge that copies all the features of the head of a maximal projection YP onto the head X that it merges with. If the phonological features of the head X are defective, X will be pronounced with the phonological features of the head Y.

NB As Conflation takes place at Merge, it is not clear that it should be defined as a PF operation. But under the assumption that lexical insertion occurs late, Conflation could be automatic while the pronunciation of a given head could be deferred until Spell Out.

Platezack to appear: Phonological information introduced in a head H can be spelled out in H or in any head of the extended projection of H.

Post-spellout operations are predicted to not interact with narrow syntax and thus to not cause effects like the link between Object Shift and verb movement in Scandinavian (Holmberg 1986).

Since both accounts are stipulative, the properties of the mechanisms postulated could easily have been different (downward movement, sideways movement...)

3. The morphology of head-movement

Head-adjunction was originally postulated to handle the fact that the probe and the goal of head-movement behave as a single constituent for all subsequent syntactic operations.

Just as important, however, is the fact that the constituent resulting from head-movement can be a locus of morphological operations (allomorphy, impoverishment, suppletion...). These operations are not defined on linear strings (Chomsky 1995:319)

3.1. Head-movement as remnant movement

Hinterhölzl 1997, Koopman and Szabolcsi 2000, Mahajan 2000, Kayne and Pollock 2001, Julien 2002, Nilsen 2003, Müller 2004, among others: head-movement is an epiphenomenon; what really happens is the movement of the maximal projection of X emptied of everything but the head X itself:
Any morphological process that may occur between L and F in this configuration would also be predicted to occur between any other two heads in the same configuration, for instance between the maximal projection in [Spec, CP] and C.

Over-generation: the Head-movement Constraint (Travis 1984) is not predicted

Further problem: the syntactic configuration for morphological processes would have to be defined recursively:

(15)

Quite apart from the blatant non-locality of this configuration, what happens if FP contains a phase (e.g., vP)?

Conclusion: we don't want morphology to operate on big chunks of syntactic structure. What we need is an operation re-bracketing a remnant movement configuration to yield a single syntactic head (Matushansky 2005).

Unless this re-bracketing operation is independently motivated, there is no reduction.

3.2. Local Dislocation

Embick and Noyer 1999, 2001, Embick 2007: the traditional head-movement approach cannot account for the prosodic constraint on synthetic comparative formation:

(16) a. smarter, #more smart
b. *beautifuller, more beautiful
Since head-movement happens before vocabulary insertion, no effect from the choice of the lexical root is expected.

Local Dislocation applies after Vocabulary Insertion and linearization, and affects both linear order and hierarchy:

(17) a. [XP X [vP [ZP Z] Y]]
b. [X * [Z * Y]]
c. [[Z Z=X] * Y]

Note that, as with Affix Hopping, a new complex head Z0 is created.

The combination of an adjective and a comparative morpheme can yield a suppled form:

(18) a. more + good → better
b. plus + bon → meilleur

Suppletion can only take place within a single head.

As Local Dislocation is a post-syntactic operation sensitive to individual lexical items, it can only apply after Vocabulary Insertion.

M-merger is defined for two structurally adjacent heads (potentially further constrained to be linearized in the same direction). Can it be redefined for the remnant movement environment in (14)?

Problem: m-merger is independently motivated (Matushansky 2006) by the need to account for cliticization in Romance and Saxon genitives (m-merger hypothesized to apply to a result of phrasal movement), as well as for suffixal definiteness marking in Danish and negative cliticization in English (m-merger hypothesized to apply without prior movement). None of these configurations are as complex as (14).

Advantages of m-merger:

▷ No need to have the same feature on the lexical head and the functional head
▷ As head-movement is a kind of syntactic movement, it targets a c-commanding position (prior to m-merger) and may yield LF effects
▷ M-merger is independently motivated

The locality of head-movement is an independent issue in this approach.

4. Movement and Reprojection

Koeneman 2000 combining Ackema, Neeleman and Weerman 1993 with Kerstens 1993: a verb enters the derivation fully inflected, then moves, re-merges and reprojects:

(20) a. Y [P] YP ZP Y
b. FP Y [P] YP ZP Y

Advantages:

▷ No need to have the same feature on the lexical head and the functional head
▷ Head-movement is a kind of movement
▷ C-command is not violated
▷ Mirror Principle (Baker 1985) is automatically derived
Koeneman 2000: motivation for head-movement is derived from some sort of requirement on predication
Shimada 2007: head-movement is a kind of QR for world and time variables

Problems:
- Romance cliticization cannot be dealt with
- "Affix Hopping"/a low position of the verb requires additional stipulations
- The treatment of auxiliaries is not clear
- Reprojection is poorly motivated independently (and saying that a head projects obligatorily whenever it merges (or re-merges) would make incorrect predictions for maximal projections that are simultaneously heads)

Summary: A beautiful idea, but probably not tenable.

5. MOVEMENT INTO A COMPLEMENT POSITION (UNDERMERGE)

Pesetsky 2010: head-movement does not result in head-adjunction; head-movement targets a complement position ("undermerge")

Independent evidence for movement into a complement position for maximal projections is scarce:
- McCloskey 1984: modern Irish has modal constructions with prepositions, which pose a problem for an obligatorily raising analysis of epistemic modals
- Sportiche 2005: some reconstruction facts suggest that the surface constituency of a determiner and a noun phrase is a result of movement
- Rosenbaum 1967, Postal 1974: Raising to Object is achieved by moving the embedded subject to the complement position of the matrix verb

Problems:
- Phrasal undermerge requires a redefinition of c-command, complementation and standard assumptions about argument structure
- Standard constituency tests become a diagnostic for a surface phenomenon

Summary: More motivation is required for a major rewrite of the framework.

6. AGREEMENT

Roberts 2010: head-movement is triggered directly by Agree, rather than by an EPP feature

Main empirical claim: head-movement can only happen when the set of features of the goal forms a subset of the set of features of the probe

Proposal: head-movement is an alternative realization of Agree:

(21) Agreement with direct objects
   a. v [prs: __, #: __] D [prs:a, #: b], [Case: __] before
   b. v [prs:a, #: b] D [prs:a, #: b], [Case: acc] after

(22) Cliticization:
   a. v [prs: __, #: __] φ [prs:a, #: b], [Case: __] before
   b. v [prs:a, #: b] (φ [prs:a, #: b]) after

Necessary ancillary assumptions:
- Romance clitics consist only of phi-features (no Case, no D, no N)
- agreement exhausts the content of the goal triggering chain reduction (Nunes 2004): the lower copy is deleted (at PF?)
- it is the higher copy that is interpreted at LF (Roberts dedicates a whole chapter to LF effects of head-movement)
- different clitics must incorporate with different functional heads (due to the lack of a Case feature)
- restructuring verbs are functional heads (cf. Cinque 2004)
- to achieve cliticization to C in some Slavic languages the complementizer and the clitics in those languages have to be assumed to contain the feature [D]

Additional ancillary assumptions:
- excorporation is possible
- a head is a phase (cf. Marantz 2001, 2006, Marvin 2002) with an edge that serves as an escape hatch
- features inside a functional head are organized hierarchically
- cliticization to T is actually cliticization to v

Problems:
- the implementation of clitic clustering is unclear; likewise for negation
- the Head-movement Constraint is not predicted
- the distinction between interpretable and uninterpretable features is blurred
- the landing site of head-movement does not c-command the extraction site (from the point of view of semantics, for sure); if it is the higher copy that is interpreted at LF, the semantic treatment of movement structures (Heim and Kratzer 1998) has to be redefined

Summary: unlikely to be correct

7. AFFIX HOPPING

Issue: How does a verb get tense morphology if it doesn't move to T?

Traditional answer: Affix Hopping (Chomsky 1957, nowadays Lowering (Emonds 1978, Embick and Noyer 1999, 2001))

(23) TP

Properties of Affix Hopping:
- has no LF effects
- is disrupted by negation (or positive verum focus)
- is not conditional on linear adjacency

Issue: how are various participles of English derived?
- A passive/perfect/continuous participle is passive/perfect/continuous without an auxiliary
- If the passive/perfect/continuous affix is located in the corresponding functional head and hops down, where are the corresponding auxiliaries inserted? (The same
question arises if participles are derived by head-movement.) If Affix Hopping is involved, why is participle formation not disrupted by (constituent) negation?

- If participles are derived by head-movement, why do English verbs move to some functional heads, but not to others?

**Issue:** do-support occurs in environments other than verum focus in declarative clauses:

(24) a. Does time fly like an arrow? question
b. Do/don't be a darling! verum focus in imperatives with he (and have)
c. Fruit flies don't like bananas, but apes do. VP-ellipsis
d. Not only did she leave me, but she also took my cat. negative inversion
e. So much did she enjoy her work, that she never rested. degree inversion
f. Shut the gate the gatekeeper certainly did. VP-topicalization

Some tentative remarks:

- For T-to-C movement in imperatives see Han and Kroch 2000
- VP-ellipsis may be correlated with VP topicalization
- Locative inversion doesn't license do-support; neither do exclamatives
- Conditional inversion is impossible without an auxiliary (Pesetsky 1989), but see Iatridou and Embick 1993
- English is unique in that negation blocks Affix Hopping

NB VP topicalization triggers do-support in German (Bader and Schmid 2006)

**Issue:** Affix Hopping violates the usually assumed bottom-up approach to derivation

The correlation between the meaning of a participle and the choice of the sister auxiliary can be achieved by Agree (if Agree can take place at Merge -- a possibility we need anyway), but sensitivity to negation cannot follow.

Tentative conclusion: However Affix Hopping is implemented, sensitivity to polarity (to an adjunct of a particular kind or the specifier of a functional head) is only necessary to account for a subset of environments for English do-support.

A new analysis of English do-support is required.

## 8. Conclusion

Desiderata for a successful theory of head-movement:

- it accounts for Romance finite verb formation, Romance cliticization and V2, at the very least
- it's a kind of movement (hence it targets a c-commanding position)
- it obeys the Head-Movement Constraint
- it is not disrupted by adjuncts or specifiers
- it allows for LF effects
- it can feed morphological processes
- whatever distinguishes it from phrasal movement is independently motivated
- it is not incompatible with periphrastic tenses or with Affix Hopping

It would be misleading (however desirable it might be) to say that head-movement is the only way of combining two syntactic terminals into one:

- Scandinavian definite articles (Hankamer and Mikkelsen 2002, 2005)
- Cuplex predicate construction (Harley, to appear)
- English verbs
- French and German preposition-determiner contractions (Cabredo Hofherr 2008)
- Dutch and German verbal clusters (Evers 1975, 2003)

It may still turn out that distinguishing head-movement from other re-bracketing phenomena is a mistake, but there is only one way to find out

### 9. References


Marantz, Alec. 2006. Phases and words. Ms., NYU.

Marvin, Tatjana. 2002. Topics in the stress and syntax of words, doctoral dissertation, MIT.


Pesetsky, David. 2010. Russian case morphology and the syntactic categories. Ms., MIT.
