Learning the A- vs Ā-movement distinction

Michael Yoshitaka Erlewine (mitcho), CNRS/Nantes

Rencontres Autour de la Linguistique Formelle, June 3–5, 2026[1]

1 Introduction

In many languages, phrasal movements fall broadly into two types:[2]

(1) **The A / Ā (A-bar) distinction:** (so-called since Chomsky 1981)

a. A-movement: feeds case/agreement, creates new antecedents for binding, more local;

b. Ā-movement: does not feed case/agreement/binding, induces crossover, allows pied-piping, licenses parasitic gaps, reconstructs more, can be long-distance.

► A remarkable fact is that there is **a good degree of cross-linguistic uniformity** in this classification, as well as which constructions fall into which class.

– Why do the various properties in (1a) vs (1b) (generally) cluster together?

– Why do cognate constructions across unrelated languages often get classed together? (I.e., why is passivization often A-movement and relativization often Ā-movement, across many unrelated languages?)

• Also note that, for some properties (e.g. binding and reconstruction possibilities), **learners (most likely) receive little to no evidence** for the type of behavior they should learn.

Today: I will introduce and explore the hypothesis that **how learners learn movement rules** can explain many of these properties.

- Learners observe regularities in their input and posit movement rules/constructions — formalized as probe specifications that obey featural Relativized Minimality — as well as hypothesized base (pre-movement) structures.
- Evidence from **properties that are (more or less) observable to the learner** — e.g. the "shape" of the movement, morphological reflexes, and distributional restrictions on the moved phrases (syntactic and semantic) — lead to the postulation of different movement rules and base structures, which lead to different movement behaviors.

Roadmap §2 Background §3 Proposal §4 Predictions and consequences §5 Conclusion

[1] For valuable discussion that has informed my thinking here, I especially thank Kenyon Branan, Isaac Gould, Matt Hewett, Hadas Kotek, Magdalena Lohninger, Ka-Fai Yip, and the audience at the Formal Linguistics in the Nordic Countries (FLINC) online colloquium.

[2] When I say "movement" today, I am referring only to phrasal movement, setting aside head movement.

2 Background: The A/Ā-distinction

(2) **Diagnostic properties of A- vs Ā-movement** (based on Safir 2019: 287)

		A	Ā	
a.	Feeding case and agreement:	+	−	
b.	Limited to nominals:	+*	−	(clauses sometimes count too)
b′.	Pied-pipes:	−	+	
c.	Crossing finite clause boundaries:	−	+	
d.	Creating new binding relationships:	+	−	
d′.	Induces "weak crossover":	−	+*	(modulo "weakest")
e.	Reconstructing for binding:	−*	+	(facts complicated)

* Some properties sometimes do *not* neatly line up, for certain constructions or languages.

See e.g. Richards 2014 and Safir 2019 for examples illustrating these properties in (2).

(3) **The positional view:** Landing sites of movement may be A-positions or Ā-positions, where an "A-position" is a "potential θ-position" (Chomsky 1981: 47).

- But see e.g. Van Urk 2015; Gong 2022; Lohninger 2024 for critiques and challenges.
- In Minimalist syntax, all movements reflect a single operation (Move, or Internal Merge). Chomsky (2004: 125) comments: "no principles can be formulated in terms of the A-/A-bar-distinction, a strong and highly controversial conclusion."

(4) **The featural view:** "the features involved in A-movement... are obligatory features of nominals, while the features that drive Ā-movement are optional" (Van Urk 2015: 26).

- But note that both the positional and featural views (3–4) do not obviously explain the behavioral differences as in (2).

(5) **The structural view:** Ā-moved-phrases are "wrapped" in additional structure that A-moved-phrases lack (with some differences, due to Horvath 2000, 2007; Cable 2007, 2010; Safir 2019; Branan and Erlewine 2023). The phrase in (b) has an adjoined "particle" (PRT).

a. A-movement of a DP:

[DP [... ___ ...

b. Ā-movement of a DP:

[[container PRT DP] [... ___ ...

- Safir (2019) proposes that the structural difference as in (5a,b) can explain a number of the distinctive properties of A- vs Ā-properties (2):
  - The added particle "'insulates' that DP from Case and agreement" (p. 285).
  - The moved DP c-commands other nominals that are crossed in (5a) but not (5b), which Safir claims explains differences in binding behavior.
  - After all, we might think that the "null theory of the A/A-bar distinction [is] presence or absence of c-command" (Müller 2026).

► However, an important question remains: **Why do some movements involve a particle and some do not?**[3]

3 Proposal

Consideration of **how learners postulate movement rules** motivates the existence of particles (PRT) that form targets of Ā-movement, as in (5b).

Some background assumptions:

- **Formal features are postulated by the learner**, based on contrasts that are meaningful in the target language (see e.g. Zeijlstra 2008; Cowper and Hall 2014; Koeneman and Zeijlstra 2014; Biberauer and Roberts 2017; Biberauer 2019).
- **Learners postulate lexically-specified movement rules** — specified as [PROBE:F] on some head — making reference to a "trigger" feature [F] (Roeper and de Villiers 2011).
- [PROBE:F] targets the closest phrase with the feature [F]. In other words, **featural Relativized Minimality** (Rizzi 1990 et seq) is innate, or reflects a general cognitive bias of dependency length minimization (see e.g. Futrell, Levy, and Gibson 2020).

3.1 Illustration

Suppose you're a language learner and you encounter many strings like this:

(6) **Basic word order:**

a. 
d. 
g. 
b. 
e.
h. 
c. 
f.
i. 

On certain limited occasions — maybe associated with some marked semantics or pragmatics — you encounter strings like this, which reflect a *marked word order*:

(7) **Construction A:**

a. 
c.
e. 
b. 
d.
f. 

- Notice that just the first word seems "out of order" in each example.

► **Let's postulate a rule** of the form [PROBE:F] = "Move the closest [F] to the front" that derives the Construction A order from a corresponding base order, which we can hypothesize.

[3] Branan and Erlewine (2023) suggest (by stipulation) that "all A'-movement is, by definition, movement of particle phrases" (p. 627). Safir strongly critiques this sort of view, saying "it would constitute a stipulation that introduces the distinction" (p. 297). Instead, he aims to derive the distribution of particles based on Case considerations (again following Rezac 2003).

(8) **Movement rule for Construction A (7):**

[PROBE: 🚗]

This rule appears to be a good hypothesis for this data because:

- The initial word in Construction A (the apparently displaced one) is always a 🚗.
- The evidence is compatible with the rule always targeting the *leftmost* 🚗:

(9) From base order ☁ 🚗 💀 🏈 🚗 (6g):

a. 🚗 [☁ __ 💀 🏈 🚗] — (7d)

b. *🚗 [☁ 🚗 💀 🏈 __] — unattested

Now suppose that, in other situations, you encounter the following:

(10) **Construction B:**

a. 🚹 🚗 🏈 ☁ 2️⃣

b. 💀 🐱 🚹 ☁

c. 🏈 💀 2️⃣ 🚗 🚹 ☁

d. 💀 🏈 ☁ 2️⃣ 🐱 🚗 🚹

e. 💀 ☁ 🚗 💀 🐱 2️⃣

f. 2️⃣ 🏈 ☁ 🚗 🚹 🐱

These too reflect a marked word order, but **apparently of a different sort of process:**

- There is **variability** in the initial (displaced) words. (It's often 💀, but not always.)
- This movement **can skip like words**, as seen from (10e): 💀 [☁ 🚗 💀 🐱 __ 2️⃣] .

► Then there must be **something that we're not seeing**. The moved words must somehow be special for the speaker. Let's hypothesize that they all have a [★] feature:

(11) **Hypothesized surface structures for (10):**

a. 🚹[★] 🚗 🏈 ☁ 2️⃣

b. 💀[★] 🐱 🚹 ☁

c. 🏈[★] 💀 2️⃣ 🚗 🚹 ☁

d. 💀[★] 🏈 ☁ 2️⃣ 🐱 🚗 🚹

e. 💀[★] ☁ 🚗 💀 🐱 2️⃣

f. 2️⃣[★] 🏈 ☁ 🚗 🚹 🐱

(12) **Corresponding hypothesized base structures for (10):**

a. 🚗 🏈 ☁ 🚹[★] 2️⃣

b. 🐱 💀[★] 🚹 ☁

c. 💀 2️⃣ 🚗 🚹 ☁ 🏈[★]

d. 🏈 ☁ 💀[★] 2️⃣ 🐱 🚗 🚹

e. ☁ 🚗 💀 🐱 💀[★] 2️⃣

f. 🏈 ☁ 🚗 2️⃣[★] 🚹 🐱

(13) **Movement rule for Construction B:**

[PROBE: ★]

The movement rule for Construction B is just as simple as that for Construction A. We are still always moving the closest match. The only difference is that we had to postulate a feature [★] which is not reflected in the surface form. [★] makes the second 💀 special in (10e/12e).

- How did [★] get there? Assume that there is an unpronounced lexical item ∅★ (possibly with semantics) that the speaker can adjoin to some phrase: [★P ∅★ 💀].

We have motivated the existence of "particles" (à la Horvath/Cable/B&E), which are adjoined to a phrase to make it a target for a particular movement, **for some movements but not others, from the process of learning movement rules.** Their postulation has the side effect of **a structural difference** for certain movement targets (5), leading to different movement behaviors as per Safir 2019.

3.2 Learning A-movement

In (many varieties of) English, **passivization must target the "closest" nominal:**[4]

(14) The bank denied [$_{DP}$ Mike] [$_{DP}$ a loan]. (Hewett 2025: 27)

a. [$_{DP}$ Mike] was denied ___ [$_{DP}$ a loan] (by the bank).

b. *[$_{DP}$ A loan] was denied [$_{DP}$ Mike] ___ (by the bank).

(15) [$_{DP}$ The kids] depended on [$_{DP}$ Krishna] for [$_{DP}$ money]. (ibid.: 31)

a. [$_{DP}$ Krishna] was depended on ___ for [$_{DP}$ money] (by the kids).

b. *[$_{DP}$ Money] was depended on [$_{DP}$ Krishna] for ___ (by the kids).

Passivization is like Construction A (7) above! Following Hewett 2025, we may propose:

(16) **Movement rule for English passives:**[5]

[PROBE:D] on T, read as "Move the closest [D] target"

Cooccurs with demotion/suppression of the agent and passive morphology.

3.3 Learning Ā-movement

In contrast, consider English *it*-cleft formation:

(17) ***It*-clefts can target various categories:**

a. It's [$_{DP}$ the keys] [that Max put ___ on the table] (, not the phone).

b. It's [$_{PP}$ on the table] [that Max put the keys ___] (, not under the sink).

c. It's [$_{DP?}$ yesterday] [that Max put the keys on the table ___] (, not today).

(18) ***It*-clefts can skip like categories:**

a. It's [$_{DP}$ Krishna] [that [$_{DP}$ the kids] depend on ___ for [$_{DP}$ money]] (, not Sarah).

b. It's [$_{DP}$ money] [that [$_{DP}$ the kids] depend on [$_{DP}$ Krishna] for ___] (, not food).

These properties suggest that **something else makes the moved phrase special**.

[4] More specifically, Hewett (2025) argues that passivization targets the *linearly leftmost*, rather than the structurally closest/highest, using independent diagnostics for "ascending" and "descending" postverbal structures.

[5] Raising could be the same, but we would have to say that experiencers of the English *seem to* type are really invisible for [PROBE:D]. See Hewett 2025 note 59 for discussion.

Clefting is like Construction B (10) above! The learner formulates a novel formal feature [CLEFT] — *the name of this feature doesn't matter!* — and postulates a lexical item (an unpronounced particle PRT) that is adjoined to these targets and projects this feature:

(19) **Hypothesized base structures for (18a,b):**

a. [$_{DP}$ the kids] depend on [$_{[CLEFT]}$ PRT [$_{DP}$ Krishna]] for [$_{DP}$ money]

b. [$_{DP}$ the kids] depend on [$_{DP}$ Krishna] for [$_{[CLEFT]}$ PRT [$_{DP}$ money]]

(20) **Movement rule for clefting:**

[PROBE:CLEFT] on a C head, read as "Move the closest [CLEFT] target"

Comes with exhaustive focus semantics and requires the *it be...* structure above.

3.4 Summary

The "shape" of a movement leads to the postulation of particles in some cases but not others.

- The postulation of particles leads to a **structural difference** (see (5a,b)) that can explain (some) Ā-movement behaviors (Safir 2019).

► Notice that this procedure also offers **a deeper explanation for Van Urk's (2015) featural description** of the A/Ā-distinction (emphasis mine):

"what distinguishes the features involved in A-movement... is that they are obligatory features of nominals, while the features that drive Ā-movement are optional" (p. 26).

- This approach also accounts for the generalization that **A-movement tends to be more local** than Ā-movement.

However, **this picture has clearly been idealized in a number of ways:**

1. There is a clear "basic" word order and the learner knows this.

   Learning movements in languages with "free word order / scrambling" or frequent pro-drop (which may be typologically linked; see e.g. Hale 1982) may look different?

2. Relevant featural distinctions amongst words and phrases in the base structures are already learned.

3. The learner can sort/cluster various non-basic word order input into separate "constructions," where different rules might apply. Other, associated morphosyntactic or semantic/pragmatic cues are likely necessary.

4 Predictions and consequences

4.1 Effects of categorial variability: Learning and losing V2

Movement to a V2 prefield is Ā-movement because (a) objects can be fronted over subjects and (b) non-nominals such as adverbs and PPs can move there too.

- ► **The categorial variability provided by (b) is particularly important,** because it forces the learner to postulate a particle.
- Meisazahl (2024) used **artificial language learning experiments**, where participants were exposed to "a recently discovered dialect of English" with V2, and concluded that **"a language containing many adjunct-initial sentences was learnt best"** (Meisazahl, Kirby, and Culbertson 2025: 1).[6]
- **This echoes the distribution of prefields in current and historical V2 languages:** "there is robust evidence that adjuncts appear very frequently in initial position in the Germanic languages" (p. 29), and similarly in Medieval Romance, but initial adjunct frequency then decreased in Romance, "support[ing] the possibility that the decline of V2 grammar is connected to the decline in adjunct-initial sentences."

4.2 Effects of semantic restrictions: Indonesian-type subjects

- ► Even for movements that do not skip other potential goals, **particles can be motivated for semantic reasons.**

- So-called "Indonesian-type" Austronesian languages exhibit a number of surface similarities, including basic "S (Aux) V O" word order, an active/passive voice alternation, no case distinctions on core arguments, and no subject φ agreement.

  Nonetheless, there is a split in the behavior of their subjects:

(21) **Properties of subjects in Indonesian-type languages:**
(synthesizing Patrianto and Chen 2023a,b; Lohninger and Katochoritis 2025)

		Javanese	Acehnese	Indonesian
a.	Must be definite/specific	+	−	−
b.	Must be nominal	??[7]	+	+
c.	Creates new antecedents	−	+	+
d.	Condition A reconstruction	+	−	−

Properties (c–d) refer to the relationship between passive subjects and agents.

- ► As subjects in Javanese are always definite/specific, the learner proposes a formal feature [α] and postulates an associated particle. We therefore predict Ā-movement behaviors (c–d) in Javanese, but not in Acehnese or Indonesian.

[6] More specifically, "participants in the adjunct-dominant condition were more likely to produce and accept sentences with novel clause-initial constituents compared to both other conditions... [and] were also better at discriminating between grammatical and ungrammatical (V3) sentences" (p. 22).

[7] Patrianto and Chen (2023b) argue that Javanese allows for PPs to occupy subject position. This too would support the postulation of a particle, but I believe such constructions to be rare.

5 Conclusion

The goal of this project is to pursue a more explanatory account of movement types and their behaviors, led by articulating **how learners would postulate movement rules and the evidence available to them** when doing so.

- ► **Properties that are more visible for the learner** (distributional evidence of the choice of movement target, its hypothesized path, and semantic restriction) lead to the postulation of **different base structures** (with and without particles) and therefore **predict different interpretational effects** (e.g. for binding and reconstruction) (Safir 2019).
- Much work remains to spell out the interpretation of different chain types, interactions with case and agreement, and further predictions of this learning-based perspective.

References

Barss, Andrew. 1986. Chains and anaphoric dependence. Doctoral Dissertation, Massachusetts Institute of Technology.

Belletti, Adriana, ed. 2004. *Structures and beyond: The cartography of syntactic structures, volume 3*. Oxford University Press.

Biberauer, Theresa. 2019. Factors 2 and 3: Towards a principled approach. *Catalan Journal of Linguistics* 45–88.

Biberauer, Theresa, and Ian Roberts. 2017. Parameter setting. In *The Cambridge handbook of historical syntax*, ed. Adam Ledgeway and Ian Roberts, 134–162. Cambridge University Press.

Branan, Kenyon, and Michael Yoshitaka Erlewine. 2023. Anti-pied-piping. *Language* 99:603–653.

Cable, Seth. 2007. The grammar of Q. Doctoral Dissertation, Massachusetts Institute of Technology.

Cable, Seth. 2010. *The grammar of Q: Q-particles, wh-movement, and pied-piping*. Oxford University Press.

Chomsky, Noam. 1981. *Lectures on government and binding*. Foris.

Chomsky, Noam. 2004. Beyond explanatory adequacy. In Belletti (2004), 104–131.

Colley, Justin, and Dmitry Privoznov. 2020. On the topic of subjects: Composite probes in Khanty. In *Proceedings of NELS 50*, ed. Mariam Asatryan, Yixiao Song, and Ayana Whitmal, 111–124.

Cowper, Elizabeth, and Daniel Currie Hall. 2014. *Reductiō ad discrīmen*: Where features come from. *Nordlyd* 41:145–164.

Futrell, Richard, Roger P. Levy, and Edward Gibson. 2020. Dependency locality as an explanatory principle for word order. *Language* 62:371–412.

Gong, Zhiyu Mia. 2022. Issues in the syntax of movement: Cross-clausal dependencies, reconstruction, and movement typology. Doctoral Dissertation, Cornell University.

Hale, Ken. 1982. Preliminary remarks on configurationality. In *Proceedings of NELS 12*.

Hewett, Matthew. 2025. The precedence component to intervention effects: Evidence from English passives. *Syntactic Theory and Research* 1:8.

Horvath, Julia. 2000. Interfaces vs. the computational system in the syntax of focus. In *Interface strategies*, ed. H. Bennis, Martin Everaert, and Eric Reuland, 183–206.

Horvath, Julia. 2007. Separating "focus movement" from focus. In *Clausal and phrasal architecture: Syntactic derivation and interpretation*, ed. Simin Karimi, Vida Samiian, and Wendy K. Wilkins, 108–145. John Benjamins.

Koeneman, Olaf, and Hedde Zeijlstra. 2014. The rich agreement hypothesis rehabilitated. *Linguistic Inquiry* 45:571–615.

Lohninger, Magdalena. 2024. Patterns in chaos: Composite A′/A probes. Doctoral Dissertation, University of Vienna.

Lohninger, Magdalena, and Ioannis Katochoritis. 2025. Between topics and subjects: an A′/A typology of Austronesian pivots. In *Proceedings of WCCFL 42*, 224–231.

Meisazahl, Marc. 2024. Learning to lose: The role of input variability in the loss of V2. Doctoral Dissertation, University of Edinburgh.

Meisazahl, Marc, Simon Kirby, and Jennifer Culbertson. 2025. Variability and learning in language change: The case of V2. *Journal of Historical Syntax* 9.

Müller, Gereon. 2026. Derivational timing and the A/A-bar distinction in syntax. Poster presented at GLOW 48, Siena.

Obata, Miki, and Samuel David Epstein. 2011. Feature-splitting internal merge: Improper movement, intervention, and the A/A′ distinction. *Syntax* 14:122–147.

Patrianto, Hero, and Victoria Chen. 2023a. A cline of Indonesian-type voice as in transition from Ā to A-syntax: Insights from four languages. Presented at AFLA 30.

Patrianto, Hero, and Victoria Chen. 2023b. Two sides of the same coin: Reappraising Indonesian-type 'pasive' and object voice in Javanese. In *Proceedings of AFLA 29*, ed. Vera Hohaus, Jens Hopperdietzel, and Siena Weingartz, 59–74.

Postal, Paul M. 1971. *Cross-over phenomena*. Holt, Rinehart, and Winston.

Rezac, Milan. 2003. The fine structure of cyclic agree. *Syntax* 6:156–182.

Richards, Norvin. 2014. A-bar movement. In *Routledge handbook of syntax*, ed. Andrew Carnie, Yosuke Sato, and Daniel Siddiqi, 167–191. Routledge.

Rizzi, Luigi. 1990. *Relativized minimality*. MIT Press.

Roeper, Tom, and Jill de Villiers. 2011. The acquisition path for *wh*-questions. In *Handbook of generative approaches to language acquisition*, ed. Jill de Villiers and Tom Roeper, 189–246. Springer.

Safir, Ken. 2019. The A-/Ā-distinction as an epiphenomenon. *Linguistic Inquiry* 50:285–336.

van Urk, Coppe. 2015. A uniform syntax for phrasal movement: A Dinka Bor case study. Doctoral Dissertation, Massachusetts Institute of Technology.

Zeijlstra, Hedde. 2008. On the syntactic flexibility of formal features. In *The limits of syntactic variation*, ed. Theresa Biberauer, 143–173. John Benjamins.