

# Arabic Nonconcatenative Morphology and the Syntax-Phonology Interface\*

Sam Zukoff

Massachusetts Institute of Technology

## 1. Introduction

This paper develops a new integrated analysis of the phonological and syntactic properties of nonconcatenative morphology in (Classical/Modern Standard) Arabic. The account centers around an algorithm for sub-word linearization at the syntax-phonology interface, here termed the *Mirror Alignment Principle* (MAP). The MAP determines the ranking of Alignment constraints (McCarthy & Prince 1993) in the phonological component based on asymmetric c-command relations in the syntax. Using the MAP, we can *predict* the exact position of all morphemes/segments in an Arabic verbal form based on their syntactic functions and structures without recourse to templates (cf. McCarthy 1979, 1981, *et seq.*).<sup>1</sup>

## 2. Puzzle

The Arabic verbal system is divided into “Forms”. These are morphosyntactic categories associated with a particular phonological shape (traditionally described in terms of a CV “template”) and a range of morphosemantics (although these have often become highly idiomatized). Within this system, Reflexive /t/ recurs across multiple Forms, sometimes as an “*infix*” (1a), sometimes as a “*prefix*” (1b).<sup>2</sup>

---

\*Special thanks to Adam Albright, Nico Baier, Kenyon Branan, Heidi Harley, Aron Hirsch, Larry Hyman, Itamar Kastner, Michael Kenstowicz, David Pesetsky, Norvin Richards, Donca Steriade, Matt Tucker, Martin Walkow, the audience at NELS 47, and audiences at Berkeley and MIT. All mistakes are my own.

<sup>1</sup>See Kastner (2016) for recent work in a similar vein on Modern Hebrew nonconcatenative morphology.

<sup>2</sup>It is unclear if it is appropriate to identify this morpheme as “Reflexive”, as it does not lead to consistent argument structure alternations typical of reflexives. All that is important is that the /t/ morpheme that shows up in multiple Forms is the exponent of the same morphosyntactic terminal (whatever that happens to be) and is in the hierarchical relations with Root that I claim it to be.

(1) *Forms with Reflexive /t/ (to example root  $\sqrt{ktb}$  ‘write’)*

a.	<b>Infixal</b>	VIII	Reflexive	( <i>ʔi</i> )k $\bar{t}$ ataba <sup>3</sup>
		V	Reflexive + Causative	$\bar{t}$ akataba
b.	<b>Prefixal</b>	VI	Reflexive + Applicative	$\bar{t}$ akaataba
		X	Causative + Reflexive	( <i>ʔi</i> )s $\bar{t}$ akataba

This distribution cannot (solely) be due to phonotactics, as the alternative affixation pattern could yield phonotactically legal structures for all categories. Form VIII could have had a phonotactically legal prefixal structure: \* $\bar{t}$ akataba, \* $\bar{t}$ akaataba. And Form V (for example) could have had a phonotactically legal infixal structure: \*kat $\bar{t}$ ataba, \*(*ʔi*)k $\bar{t}$ ataba. Something more complicated must be involved in determining this distribution.

### 3. Proposal: The Mirror Alignment Principle

Most previous accounts of these facts have had to stipulate special behavior of Form VIII.<sup>4</sup> McCarthy (1981) posits a morpheme-specific autosegmental reassociation rule (“Eighth Binyan [= Form] flop”). Ussishkin (2003, *et seq.*) claims that the /t/ is not the same morpheme across Forms, and needs to posit different Alignment rankings with respect to ALIGN-ROOT for different morphemes. Tucker (2010) follows the single /t/ morpheme approach, but must posit that, when this morpheme appears in Form VIII, it is indexed to a special Alignment constraint, ranked differently with respect to ALIGN-ROOT. There is, however, a *syntactic generalization* about this (morpho)phonological distribution of the Reflexive /t/ that these analyses seem to have missed:

- (2) a. When Reflexive co-occurs with (and scopes over) another verbal derivational morpheme, e.g. Causative or Applicative (cf. (14–15)), it is *prefixal*.
- b. When it is the only verbal derivational morpheme, it is *infixal*.

If we can directly relate syntactic structure to phonological behavior, then we can use this generalization to account for the apparent idiosyncrasy of the Reflexive. I implement this generalization within a new approach to *sub-word linearization*: the crux of the proposal is an interface algorithm that translates hierarchical syntactic relations into phonologically-interpretable information (i.e. rankings of ALIGNMENT constraints). This algorithm, which I term the *Mirror Alignment Principle (MAP)*, is defined in (3). (See Zukoff 2017 for development of this approach for morpheme-ordering problems in the Bantu languages.)

<sup>3</sup> *ʔi* is epenthesized when the word is initial within a prosodic phrase, to avoid phrase-initial clusters.

<sup>4</sup>For other recent accounts, see additionally Tucker (2011), Wallace (2013).

## Arabic Nonconcatenative Morphology

(3) *The Mirror Alignment Principle (MAP)*

If terminal node  $\alpha$  asymmetrically c-commands terminal node  $\beta$  (in the output of the syntactic/morphological component), then ALIGN- $\alpha$  dominates ALIGN- $\beta$  (in the phonological component).

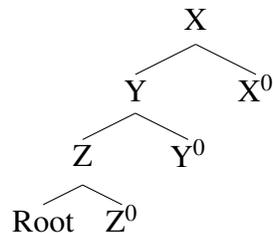
The MAP governs the ranking of classic gradient Alignment constraints (McCarthy & Prince 1993), defined schematically in (4).

(4) ALIGN- $\alpha$ -LEFT/RIGHT:

Assign a violation for each segment that intervenes between the Left/Right edge of (the phonological exponent of)  $\alpha$  and the Left/Right edge of the word.

In this proposal, linearization is enacted in an Optimality Theoretic (Prince & Smolensky 2004) phonological component *by Alignment*. Morphology provides an unordered set of morphemes for the phonological input. The MAP provides a ranking of ALIGNMENT constraints in CON based on the syntactic structure. GEN produces a candidate set consisting of all possible morpheme orders. EVAL selects the output candidate which is most harmonic with respect to CON, i.e. the ordered ranking of ALIGNMENT constraints, FAITHFULNESS constraints, and MARKEDNESS constraints. This system is illustrated in (5–7):

(5) *Schematic Example: complex head (result of head movement)*



(6) *C-Command Relations & MAP-determined Ranking*

- a.  $X^0$  asymmetrically c-commands  $Y^0$ ,  $Z^0$ , and Root  
 $\Rightarrow$  ALIGN-X  $\gg$  ALIGN-Y, ALIGN-Z, ALIGN-ROOT
- b.  $Y^0$  asymmetrically c-commands  $Z^0$  and Root  
 $\Rightarrow$  ALIGN-Y  $\gg$  ALIGN-Z, ALIGN-ROOT
- c.  $Z^0$  and Root *symmetrically* c-command each other  
 $\Rightarrow$  No MAP-determined ranking between ALIGN-Z and ALIGN-ROOT
- d. **Total ranking:**  
**ALIGN-X  $\gg$  ALIGN-Y  $\gg$  ALIGN-Z, ALIGN-ROOT**

(7) *Tableau for schematic derivation*

/X, Y, Z, ROOT/	ALIGN-X-L	ALIGN-Y-L	ALIGN-Z-L	ALIGN-ROOT-L
☞ a. X-Y-Z-ROOT		*	**	***
☞ b. X-Y-ROOT-Z		*	***	**
c. X-Z-Y-ROOT		**!	*	***
d. Y-X-Z-ROOT	*!		**	***

The tableau in (7) shows how the Alignment ranking generated by the MAP in (6) for the tree in (5) selects a morpheme order for that derivation.<sup>5</sup> Since  $X^0$  is the highest terminal in the tree and thus c-commands all the other terminals, the MAP ranks ALIGN-X-L highest. This rules out any order that does not place X at the left edge, here represented by (7d). The next highest terminal is  $Y^0$ , and thus ALIGN-Y-L comes next in the ranking. Among all remaining candidate orders (i.e. those with X at the left edge), this eliminates any which does not have Y immediately following X, here represented by (7c).<sup>6</sup>

The MAP alone cannot adjudicate between the remaining candidate orders (7a) and (7b). Language-specific factors (e.g. default rankings) will have to apply to resolve under-determined rankings like that of ALIGN-Z-L and ALIGN-ROOT-L where necessary. Arabic employs a specific strategy (see (12)) that is applicable across the system.

This has demonstrated that the MAP allows us to predict the position of all segments in an Arabic verbal form, including infixes and peripheral affixes, based on their syntactic functions/structures, in conjunction with phonotactics and other phonological considerations. In the remaining sections, I will illustrate how this framework derives the infix vs. prefix distinction for different types of Reflexives and Causatives. In general, this approach allows for an integrated syntactic and phonological analysis of the entirety of the Arabic verbal system (see (14–15)).

#### 4. Analysis of Reflexive

An Alignment-based analysis of the Reflexive requires an apparent ranking paradox (cf. Tucker 2010), as shown in (8). That these rankings properly derive the distribution is confirmed in (9) and (10). (Alignment violations are not tallied for the epenthetic segments *?i*; counting these violations would not change the evaluations.)

(8) *Ranking paradox*

- a. Prefixal Forms (V,VI,X): ALIGN-REFLEXIVE-L  $\ggg$  ALIGN-ROOT-L
- b. Infixal Form (VIII): ALIGN-ROOT-L  $\ggg$  ALIGN-REFLEXIVE-L

<sup>5</sup>In this illustration, all morphemes are given left-oriented Alignment constraints, and Alignment violations are assigned as if each morpheme is a single segment.

<sup>6</sup>This demonstrates why the Alignment constraints must be defined gradiently in this system. If they were defined categorically (cf. McCarthy 2003), ALIGN-Y-L would not be able to differentiate between the candidates which displace Y from the left edge.

*Arabic Nonconcatenative Morphology*

(9) *Alignment Derivation of Form V*

/t, $\mu_c$ , ktb, a, a/	ALIGN-REFL-L	ALIGN-ROOT-L
☞ a. <u>t</u> ak <sub>c</sub> ataba		**
b. (?i)k <u>t</u> at <sub>c</sub> ataba	*!	

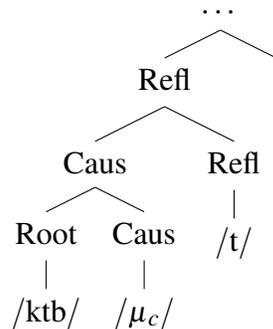
(10) *Alignment Derivation of Form VIII*

/t, ktb, a, a/	ALIGN-ROOT-L	ALIGN-REFL-L
a. <u>t</u> aktaba	*!	
☞ b. (?i)k <u>t</u> ataba		*

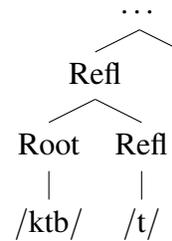
The MAP provides a solution for the apparent paradox. The two types have different syntactic structures (shown in (11)). Therefore, the MAP generates *distinct Alignment rankings* (as required in (9) and (10)). Note that this requires that Alignment rankings be able to differ across phonological derivations, contrary to typical conceptions of OT. But this is not a bug in the system – this is the basis on which system operates.

(11) *Syntactic Structures with Reflexive*

a. **Form V** *takat<sub>c</sub>ataba*



b. **Form VIII** (?i)kataba



In Form V, Refl *asymmetrically c-commands* Root. Therefore, the MAP generates the ranking ALIGN-REFL-L  $\gg$  ALIGN-ROOT-L, which is required to derive the prefixal behavior of /t/, as in (9). In Form VIII, on the other hand, Refl and Root stand in *symmetric c-command*. Since the MAP only asserts rankings based on asymmetric c-command, the ranking between ALIGN-REFL-L and ALIGN-ROOT-L is underdetermined. This requires one stipulation: across the board in Arabic, ranking indeterminacy caused by absence of asymmetric c-command is resolved by the principle in (12):

(12) *Default ranking statement for Arabic:*

When the MAP provides no ranking statement (i.e. when two heads are not in asymmetric c-command), ALIGN-ROOT-L is top-ranked by default.

This resolves the ranking for Form VIII as ALIGN-ROOT-L  $\gg$  ALIGN-REFL-L, which is required to derive the infixal behavior of /t/, as in (10). The application of this default

ranking in cases of indeterminacy accounts for infixal behavior across the system, including in Form II below.

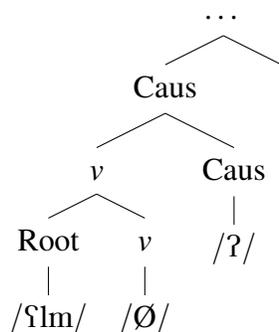
## 5. Analysis of Causative

Arabic has two types of basic causatives (cf. Wright 1896, *a.o.*). The first is Form II, which is marked by an infixal consonantal mora ( $/\mu_c/$ ), and has a fairly wide range of transitivizing semantics, including causative. The other is Form IV, which is marked by a prefixal  $/ʔ/$ , and has fairly consistent causative semantics. The root  $\sqrt{\text{ʕlm}}$  ‘know’ provides us with an ideal minimal pair. It has a Form II causative  $\text{ʕal}_c\text{lama}$  which means ‘teach’, and it also has a Form IV causative  $\text{ʔaʕlama}$ , which means ‘inform’ ( $\approx$  ‘make s.o. know’).

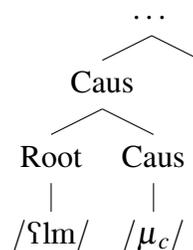
The syntax in (13), which treats Form II as a root-selecting causative and Form IV as a  $v$ P-selecting causative (where  $v$  has a null exponent), captures both the semantic properties and the ordering properties:

### (13) Syntactic Structures with Causative

#### a. Form IV $\text{ʔaʕlama}$



#### b. Form II $\text{ʕal}_c\text{lama}$



On the semantics side, we would expect a root-selecting head to allow more idiomatic semantics than a non-root-selecting head (Marantz 1997). The root-selecting Caus head in Form II yields a wide range of semantics, as expected. The  $v$ P-selecting Caus head in Form IV yields consistently causative semantics, as expected.

On the ordering side, the syntactic distinction creates an ordering distinction via the MAP. In Form IV, Caus asymmetrically *c*-commands Root. The MAP thus generates the prefixal ranking ALIGN-CAUS-L  $\gg$  ALIGN-ROOT-L. On the other hand, in Form II, Caus and Root are in symmetric *c*-command, so the MAP provides no ranking. The default ranking statement in (12) applies, generating the infixal ranking ALIGN-ROOT-L  $\gg$  ALIGN-CAUS-L.

## 6. Summary of Verbal System

The analysis of the remaining verbal Forms is outlined in (14) and (15). The syntactic analyses posited here, coupled with a few additional morphophonological constraints and assumptions (e.g. Reflexive  $/t/$  and Causative  $/ʔ/$  must surface in prevocalic position), derive the full range of phonological structures in the core of the verbal system.

*Arabic Nonconcatenative Morphology*

(14) *Morphosyntactic structure of verbal Forms*

Form	Perf. Act.	Syntactic structure	Alignment Ranking
I	<i>kataba</i>	[v [Root]]	(ALIGN-RT-L ≫ ALIGN-v-L)
II	<i>kat<sub>c</sub>taba</i>	[Caus [Root]]	ALIGN-RT-L ≫ ALIGN-CAUS-L
III	<i>kaa<sub>v</sub>taba</i>	[Appl [Root]]	ALIGN-RT-L ≫ ALIGN-APPL-L
IV	<i>ʔaktaba</i>	[Caus [v [Root]]]	ALIGN-CAUS-L ≫ ALIGN-RT-L (≫ ALIGN-v-L)
V	<i>takat<sub>c</sub>taba</i>	[Refl [Caus [Root]]]	ALIGN-REFL-L ≫ ALIGN-RT-L ≫ ALIGN-CAUS-L
VI	<i>takaa<sub>v</sub>taba</i>	[Refl [Appl [Root]]]	ALIGN-REFL-L ≫ ALIGN-RT-L ≫ ALIGN-APPL-L
VII	<i>(ʔi)n<sub>k</sub>kataba</i>	[Mid [v [Root]]]	ALIGN-MID-L ≫ ALIGN-RT-L (≫ ALIGN-v-L)
VIII	<i>(ʔi)kt<sub>a</sub>taba</i>	[Refl [Root]]	ALIGN-RT-L ≫ ALIGN-REFL-L
X	<i>(ʔi)st<sub>a</sub>ktaba</i>	[Caus [Refl [v [Root]]]]	ALIGN-CAUS-L ≫ ALIGN-REFL-L ≫ ALIGN-RT-L

(15) *Morphemes involved in verbal Forms*

Syntactic Heads	Morphs	Forms
Applicative	/μ <sub>v</sub> /	III, VI
Reflexive	/t/	V, VI, VIII, X
Middle	/n/	VII
v	/∅/	I, IV, VII, X
Causative	i. /μ <sub>c</sub> / (sister to Root)	II, V
	ii. /ʔ/~s/ (elsewhere)	IV, X

**7. Conclusion**

The MAP approach offers new insights about the relationship between the verbal (morpho)syntax of Arabic and its (morpho)phonological system, and provides a more complete and consistent account of its phonological complexities and typological unusualness. Adopting the MAP approach also brings nonconcatenative morphological processes under the umbrella of phenomena which can illustrate the Mirror Principle:

(16) *The Mirror Principle* (Baker 1985:375)

“Morphological derivations must directly reflect syntactic derivations (and vice versa).”

By using *Alignment rankings* determined via phonological analysis, rather than just linear order, to infer the underlying word-internal structure, we can apply Mirror Principle reasoning to infer syntactic structure from surface morpheme order for any sort of morphological system, concatenative or otherwise.

Sam Zukoff

## References

- Baker, Mark. 1985. The Mirror Principle and Morphosyntactic Explanation. *Linguistic Inquiry* 16:373–415.
- Kastner, Itamar. 2016. Form and Meaning in the Hebrew Verb. Doctoral dissertation, NYU.
- Marantz, Alec. 1997. No Escape from Syntax: Don't Try Morphological Analysis in the Privacy of Your Own Lexicon. *University of Pennsylvania Working Papers in Linguistics* 4:14.
- McCarthy, John J. 1979. Formal Problems in Semitic Phonology and Morphology. Doctoral dissertation, MIT.
- McCarthy, John J. 1981. A Prosodic Theory of Nonconcatenative Morphology. *Linguistic Inquiry* 12:373–418.
- McCarthy, John J. 2003. OT Constraints are Categorical. *Phonology* 20:75–138.
- McCarthy, John J., & Alan Prince. 1993. Generalized Alignment. In *Yearbook of Morphology 1993*, ed. Geert Booij & Jaap van Marle, 79–153. Kluwer.
- Prince, Alan, & Paul Smolensky. 2004. *Optimality Theory: Constraint Interaction in Generative Grammar*. Malden, MA: Blackwell Publishing.
- Tucker, Matthew A. 2010. Roots and Prosody: The Iraqi Arabic Derivational Verb. *Recherches linguistiques de Vincennes* 39:31–68.
- Tucker, Matthew A. 2011. The Morphosyntax of the Arabic Verb: Toward a Unified Syntax-Prosody. In *Morphology at Santa Cruz: Papers in Honor of Jorge Hankamer*. University of California, Santa Cruz: Linguistics Research Center.
- Ussishkin, Adam. 2003. Templatic Effects as Fixed Prosody: The Verbal System in Semitic. In *Research in Afroasiatic Grammar II*, ed. Jacqueline Lecarme, 511–530. Amsterdam: John Benjamins Publishing.
- Wallace, Katherine. 2013. A Concatenative Approach to Semitic Templatic Morphology. Ms., NYU.
- Wright, William. 1896. *A Grammar of the Arabic Language*, volume I. Cambridge: Cambridge University Press, 3rd edition.
- Zukoff, Sam. 2017. The Mirror Alignment Principle: Morpheme Ordering at the Morphosyntax-Phonology Interface. In *Papers on Morphology: MIT Working Papers in Linguistics 81*, ed. Snezana Iovtcheva & Benjamin Storme, 105–124. Cambridge, MA: MITWPL.

Sam Zukoff

szukoff@mit.edu, samzukoff@gmail.com