

## À propos d'une alternance de longueur vocalique en arabe (ou, «pourquoi la *base* est pauvre... »)

1. Dans la majorité des dialectes de l'arabe, toute voyelle est réalisée *brève* en position finale de mot, et *longue* lorsqu'elle est suivie d'un suffixe.

<i>arabe du Caire :</i>	ábú      père	abú:ja      mon père
	kúnti      tu étais (fém.)	ma kuntí:f      tu n'étais pas (fém.)
	ʔúlu      dites!	ʔulú:li      dites-moi!

<i>arabe de Tunis :</i>	ǧárbu      ils ont frappé	ǧarbú:ni      ils m'ont frappé
	mʔá      avec	mʔá:ja      avec moi
	ʔándi      j'ai (= chez moi)	ma ʔandí:f      je n'ai pas

2. Une représentation sous-jacente? Quelle représentation sous-jacente?

Dans le cadre de la phonologie générative, les diverses solutions imaginables ont été imaginées...

- Une représentation sous-jacente brève et une règle d'allongement devant suffixe (Bro-selow, 1976).
- Une représentation sous-jacente longue et une règle d'abrègement en position finale de mot (Abdo, 1969)
- Des représentations sous-jacentes longues et brèves, accompagnées des règles correspondantes (Angoujard, 1979)

3. Une proposition récente :

Soucieux de défendre l'existence de représentations sous-jacentes dans le cadre de la Théorie de l'Optimalité (*contra* Burzio, 1996), McCarthy (2003) a rejoint le camp des générativistes favorables à une dérivation du type /abu:/ → [ábu].

4. Objectif de cette communication :

- Proposer une analyse monostratale de ces alternances, dans le cadre de la Phonologie Déclarative (Bird, 1990, 1995; Angoujard, 1997a).

Rappel: Un niveau de représentation *unique*. Toute représentation est la somme de représentations partielles. Les représentations partielles sont des *contraintes*. Ces contraintes sont, à l'évidence, non ordonnées et d'application obligatoire. Angoujard (1997a) propose une intégration de ces contraintes dans une analyse par «principes et paramètres».

- Montrer que cette alternance peut donc être appréhendée en l'absence de toute référence à une quelconque représentation sous-jacente (et, naturellement, à la «richesse» d'une *base* excluant toute restriction au sein des représentations sous-jacentes potentielles).

## 5. Remarque préliminaire :

Dans l'approche de cette alternance de longueur vocalique, la présence (et le rôle) de l'*accent de mot* est généralement ignoré. Il ne joue, en particulier, aucun rôle dans l'analyse de McCarthy (2003). Il apparaissait, au mieux, dans les travaux des années 70, au travers d'une règle crucialement ordonnée avant ou après les règles d'abrègement/allongement.

Or,

- les représentations sont *unes*, qui comportent des voyelles brèves et longues, accentuées et inaccentuées (peut-être également, nous y reviendrons, tendues et lâches) ;
- c'est bien cet ensemble qui est en cause ici, c'est bien la *prosodie*.

## 6. Accent de mot et longueur vocalique dans l'arabe du Caire (Mitchell, 1962).

Pourquoi retenir l'arabe du Caire ?

- C'est cette variété qui est analysée dans McCarthy (2003).
- C'est une accentuation plutôt spécifique, assez complexe et qui a fait couler beaucoup d'encre (McCarthy, 1979; Kenstowicz, 1979; Broselow, 1982; Halle and Vergnaud, 1987; Angoujard, 1990).

6.1. Si la dernière syllabe est de type [C $\bar{V}$ C] ou [CVCC], l'accent est final: [katabúh] «ils l'ont écrit», [fihímt] «j'ai compris».

6.2. Si la dernière syllabe est de type [CV] ou [CVC], alors :

- Si l'avant-dernière est de type [CV], et si l'antépénultième est de type [CV], alors la syllabe antépénultième est accentuée: [kátabu] «ils ont écrit», [kátabit] «elle a écrit».
- Si l'avant-dernière est de type [Ci] ou [Cu], et si l'antépénultième est de type [C $\bar{V}$ ], alors la syllabe antépénultième est accentuée. Cette accentuation est souvent «cachée» parce que la voyelle pénultième n'est pas réalisée... [sáf(i)ru] «ils ont voyagé».

6.3. Dans tous les autres cas, l'accent est pénultième: [muʔállim] «instituteur», [maknása] «balai».

N.B. Les «sur-lourdes» finales sont interprétées comme des syllabes lourdes suivies d'une syllabe légère à noyau vide, cf. l'accentuation de la forme [ma kuntí:f].

6.4. On ne rencontre de voyelles longues que sous l'accent (*i.e.* toute voyelle longue est accentuée).

6.5. Il existe une alternance [V] ~ [V̄] régulière («*fin de mot*» ~ «*devant suffixe*»). Cette alternance est strictement limitée à la voyelle précédant le *dernier* suffixe (elle est classiquement analysée comme «cyclique») :

[kátabu], [katabúha], [makatabuháf]

Remarques :

Les points 6.4. et 6.5. méritent d'être commentés.

- Le fait que la longueur vocalique soit neutralisée, sauf sous l'accent de mot, est très commun (c'est également vrai à Tunis et, à ma connaissance, dans les autres dialectes).

- Cette neutralisation interdit de simplement supposer qu’une voyelle est longue devant suffixe... la marque du pluriel n’est pas réalisée longue (mais peut-être «tendue», cf. Mitchell, 1962) dans [makatabuháʃ].
7. Une analyse de l’accentuation dans le dialecte du Caire (Angoujard, 1990)
- 7.1. Deux types de syllabes sont reconnus :
- syllabe légère (L) = [CV] (également [C\*])
  - syllabe lourde (H) = [CVC] et [C $\bar{V}$ ] (également [CVC\*]).
- 7.2. La syllabe finale est toujours *interprétée* comme H.
- 7.3. Les pieds binaires sont de type [+–].
- 7.4. Seuls sont construits (de gauche à droite) les pieds binaires [L L] et [[C $\bar{V}$ ] [Cv]] (où v = [i] ou [u]).
- 7.5. Comme pour tout dialecte de l’arabe, l’accent de mot porte sur l’avant-dernier pied (sur la partie forte du pied, s’il s’agit d’un pied binaire).

8. Interprétation *via* le modèle rythmique (Angoujard, 1997b) et sous forme de contraintes :

$$C\_mod: \quad mod = mod\_L \vee mod\_H$$

«Un modèle rythmique est soit “léger” (L), soit “lourd” (H)»

$$C\_mod\_L: \quad mod\_L = \left\{ init \wedge som \right\}_{\square} \wedge \neg(\square ]\omega)$$

«Un modèle L est de type [1 2] et il ne peut pas se trouver en fin de mot»

$$C\_mod\_H: \quad mod\_H = \left\{ \begin{array}{l} \text{a): } \left\{ init \wedge som \wedge fin \right\} \wedge \neg(\square ]\omega) \\ \vee \\ \text{b): } \left\{ mod\_L \wedge (mod\_L \mid som \mid Seg_{\square} \wedge var(\square)) \right\} \\ \vee \\ \text{c): } \left\{ init \wedge som \right\}_{\square} \wedge (\square ]\omega) \end{array} \right\}$$

«Un modèle H est soit de type [123] (dans ce cas on ne peut pas le trouver en fin de mot), soit de type [121\*], soit de type [12] (seulement lorsqu’il est en fin de mot)»

#### Commentaires :

- La partie a) de la contrainte C\_mod\_H exclut tout modèle à trois positions en fin de mot. Ceci a pour conséquence que dans une forme comme [kal] «il a mangé» (avec un [a] *bref*) la liquide finale ne peut être interprétée comme une position 3. Le fait que la dernière consonne peut toujours recevoir un affixe vocalique (cf. [kálit] «elle a mangé») va dans ce sens. Par contre, ce choix contredit le principe (proposé ailleurs, cf. Angoujard, 2004) de *minimalité*. . . Une solution serait de restreindre l’interdiction au seul cas où *som* et *fin* ont la même valeur pour SEG.

- Les contraintes C<sub>mod</sub> et C<sub>mod\_H b</sub>) appartiennent à GU.
- Les contraintes C<sub>mod\_L</sub> et C<sub>mod\_H a</sub>) et c) (qui font référence à la frontière de mot) sont apparemment spécifiques à la langue considérée. GU retient :

$$\begin{array}{l}
 \text{C}_{\text{mod}_L}(\text{GU}): \quad \text{mod}_L = \left\{ \text{init} \wedge \text{som} \right\}_{\mathbb{I}} \\
 \\
 \text{C}_{\text{mod}_H}(\text{GU}): \quad \text{mod}_H = \left\{ \begin{array}{l}
 \text{a): } \left\{ \text{init} \wedge \text{som} \wedge \text{fin} \right\} \\
 \vee \\
 \text{b): } \left\{ \text{mod}_L \wedge \left( \text{mod}_L \mid \text{som} \mid \text{Seg}_{\mathbb{I}} \wedge \text{var}(\mathbb{I}) \right) \right\}
 \end{array} \right\}
 \end{array}$$

- On peut toutefois supposer que les contraintes C<sub>mod\_L</sub> et les parties a) et c) de C<sub>mod\_H</sub> sont communes à toutes les langues dans lesquelles l'opposition  $[\check{V}]/[\bar{V}]$  est pertinente (c'est au moins vrai pour le latin).
- La «curiosité» que constitue la partie c) de C<sub>mod\_H</sub> pourrait être reliée à une notion de «voyelle moyenne» (neutralisation de la distinction de longueur) et/ou à sa caractérisation acoustique comme voyelle *tendue* (cf. le conseil donné dans Mitchell, 1962 : «Pronounce ii and final i approximately as the vowel in English 'beet'»).

9. Que manque-t-il?

Contrainte C<sub>δ<sub>ω</sub></sub> : le dernier modèle rythmique du dernier *domaine* avant la frontière de mot est obligatoirement de type [123] *mod\_H*.

Soit il s'agit d'une instance de *mod\_H b*) (par exemple, dans [katabitha]), soit, *par défaut*, d'une instance de *mod\_H a*).

10. Contraintes sur la construction des pieds (valeurs paramétriques pour l'arabe du Caire) :

10.1.  $\Sigma = \Sigma_{un} \vee \Sigma_{bin\_G}$

10.2.  $\Sigma_{un} \equiv \Sigma[[\text{mod}]_{[1]} \wedge ([1]|\text{VAL-}\Sigma: +)$

10.3.  $\Sigma_{bin\_G} \equiv (\Sigma[[\text{mod}]_{[1]} \wp [\text{mod}]_{[2]}\Sigma)$   
 $\wedge ([1] = (\text{L} \vee \text{H}(\bar{v}))$   
 $\wedge (([1] = \text{H}(\bar{v})) \Rightarrow ([2] = \text{L}(i/u)))$   
 $\wedge ([1]|\text{VAL-}\Sigma: +) \wedge ([2]|\text{VAL-}\Sigma: -)$

10.4. *max*<sub>Σ<sub>bin</sub></sub> (GU) : le plus grand nombre possible de pieds binaires sont construits.

11. Place de l'accent de mot en arabe :

$$\left[ \dots \Sigma | \text{mod} \begin{bmatrix} \text{POIDS-}\Sigma: + \\ \text{POIDS-}\omega: 4 \end{bmatrix} \wp \Sigma | \text{mod} \begin{bmatrix} \text{POIDS-}\Sigma: + \\ \text{POIDS-}\omega: 3 \end{bmatrix} \right]_{\omega}$$

12. La forme [katabúha]

Représentations lexicales :

- Radical : <k, a, t, a, b>
- Suffixe\_1 : <u>
- Suffixe\_2 : <h, a>

Soit, après concaténation :

$$\omega [ \delta [k, a, t, a, b] \delta [u] \delta [h, a] ] \omega$$

ou, peut-être, si la concaténation morphologique est cyclique :

$$\omega [ \delta [ \delta [k, a, t, a, b] \delta [u] \delta ] \delta [h, a] ] \omega$$

Autres contraintes :

$$\phi \rightarrow mod^+$$

$$(C\_mod \wedge C\_mod\_L \wedge C\_mod\_H) \Rightarrow$$

- <h, a> = {init  $\wedge$  som} (C\\_mod\\_H b)
- <b, u> = {init  $\wedge$  som  $\wedge$  fin} (C\\_ $\delta$ \\_ $\omega$ )
- <t, a> = {init  $\wedge$  som} (C\\_mod\\_L)
- <k, a> = {init  $\wedge$  som} (C\\_mod\\_L)

$$\Sigma \wedge \Sigma_{un} \wedge \Sigma_{bin\_G} \Rightarrow$$

- <h, a> = [VAL\_ $\Sigma$ : +]  $\wedge$  [POIDS\_ $\omega$ : 3]
- <b, u> = [VAL\_ $\Sigma$ : +]  $\wedge$  [POIDS\_ $\omega$ : 4]
- <t, a> = [VAL\_ $\Sigma$ : -]  $\wedge$  [POIDS\_ $\omega$ : 2]
- <k, a> = [VAL\_ $\Sigma$ : +]  $\wedge$  [POIDS\_ $\omega$ : 3]

	x		x		x		x		x
	x	x	x	x	x	x	x	x	x
	k	a	t	a	b	u	h	a	

13. Représentation (simplifiée) du modèle <bū> après *unification* :

$$mod\_H \left[ \begin{array}{c} \left[ \begin{array}{l} \text{SEG: b} \\ \text{POS: 1} \\ \text{PC: } i \end{array} \right]_{init} \left[ \begin{array}{l} \text{SEG: u} \boxed{1} \\ \text{POS: 2} \\ \text{PC: } s \boxed{2} \end{array} \right]_{som} \left[ \begin{array}{l} \text{SEG: } \boxed{1} \\ \text{POS: 3} \\ \text{PC: } \boxed{2} \end{array} \right]_{fin} \end{array} \right]$$

Question : Comment éviter une unification conduisant à \*[katabúhha]?

Cf. les réalisations diverses des formes 3F. Sg. de l'accompli + pronom :

- [ḍarbítu] (arabe du Caire)
- [ḍarbittu] (arabe de Tunis)
- [ḍarbáto] (Algérie)

#### 14. Conclusion

Cette alternance de longueur vocalique peut parfaitement être analysée pour ce qu'elle est : une alternance. . .

Ceci rend parfaitement inutile tout recours à une «base» constituée d'un ensemble potentiellement infini de séquences de segments.

Ceci rend également inutile tout recours à des représentations sous-jacentes (et donc à des règles ou contraintes destructrices ou modificatrices).

Les formes attestées (plus précisément, les *représentations* attestées, *i.e.* les formes *interprétées*) correspondent seulement aux représentations lexicales (pour l'essentiel, des séquences ordonnées de segments) associées (*unifiées*) aux contraintes issues de l'interprétation *prosodique*.

## Références

- Daud A. Abdo. *On Stress and Arabic Phonology: A Generative Approach*. khayats, Beirut, 1969.
- Jean-Pierre Angoujard. Allongement et abrégement vocalique en arabe tunisien. *Analyses, Théorie*, 1979/1:1–12, 1979.
- Jean-Pierre Angoujard. *Metrical Structure of Arabic*. Foris, Dordrecht, 1990. 304 p.
- Jean-Pierre Angoujard. La phonologie déclarative. *Langages*, 125:35–54, 1997a.
- Jean-Pierre Angoujard. *Théorie de la Syllabe. Rythme et Qualité*. CNRS Editions, Paris, 1997b. 224 p.
- Jean-Pierre Angoujard. Structure syllabique et information lexicale en arabe. In *Journées d'Etudes sur la Parole*, 2004. Fès, 19-22 avril.
- Steven Bird. *Constraint-Based Phonology*. PhD thesis, University of Edinburgh, 1990.
- Steven Bird. *Computational Phonology. A constraint-based approach*. Cambridge University Press, Cambridge, 1995. 203 p.
- E. Broselow. *The Phonology of Egyptian Arabic*. PhD thesis, University of Massachusetts, 1976.
- E. Broselow. On predicting the interaction of stress and epenthesis. *Glossa*, 16:115–132, 1982.
- Luigi Burzio. Surface constraints versus underlying representations. In J. Durand and B. Laks, editors, *Current Trends in Phonology: Models and Methods*, pages 123–142. European Studies Research Institute, University of Salford, 1996.
- Morris Halle and Jean-Roger Vergnaud. *An Essay on Stress*. M.I.T. Press, Cambridge, Massachusetts, 1987.
- Michael Kenstowicz. Notes on Cairene Arabic syncope. In *Proceedings of the 10th Annual Conference on African Linguistics*, 1979.
- John J. McCarthy. On stress and syllabification. *Linguistic Inquiry*, 10(3), 1979.
- John J. McCarthy. Richness of the base and the determination of underlying representations. [disponible à ROA], 2003.
- T. F. Mitchell. *Colloquial Arabic*. Hodder & Stoughton, 1962.

Merci à  et à L<sup>A</sup>T<sub>E</sub>X