Characterising the Arabic Sound System
(Consonant Resonance and Phonological Representations)

Alex Bellem, SOAS[1]
a.bellem@cbrl.org.uk

Most broadly, this paper is concerned with the issue of sound system typology, focusing on resonance in consonants (namely, the elements I, U and A), and the application of this to Arabic. I shall argue that the phenomenon of 'emphasis' in Arabic is a by-product of the fundamental patterning of the sound system into resonance characteristics, and thus not a phenomenon to be seen in isolation. I argue further that the non-emphatic coronals of Arabic are 'front' consonants – characterised by the element I – and that Arabic words consist of resonance domains. Lastly, while emphatic consonants (among others) are characterised by the element A, not all instances of what is often termed 'emphasis' involve A.

1 Background

- Arabic emphatic consonants – usually said to be characterised predominantly by secondary pharyngealisation / velarisation.
- 'Primary' emphatics = a set of 'pharyngealised' coronal obstruents (*ṭ ṣ ḍ ~ ð̣*, depending on the dialect)[2]; typically said to trigger 'emphasis spread'.
- Consonants susceptible to 'emphasis' – 'secondary' emphatics (typically at least *ṛ ḷ ṃ ḅ*).
- 'Emphasis spread' analysed for some dialects as unbounded leftwards but blockable rightwards (typically by palatals, depending on the dialect).[3]

1.1 Problematic data *1

- Many dialects have lexemes said to be at least partly emphatic, but with no 'primary' emphatic consonant. Typical minimal pairs often cited for various dialects:[4]

(1a)	*ḅāḅa*	'Daddy'	(1b)	*bāb-a*	'his door'
	walḷā(h)	'by God'		*walla*	'he appointed'
	ṃayy	'water'		*mayyit*	'dead (m.s.)'
	ǰāṛ-i	'my neighbour (m.s.)'		*ǰāri*	'flowing (m.s.)'

- In (1a), the 'emphatic' examples always involve the low vowel (*a / ā*); in (1b), the low vowel *per se* does not trigger 'emphasis'.

1.2 Problematic data *2

- In some dialects, 'emphaticness' seems to arise from certain consonant combinations, e.g. Muslim Baghdadi Arabic:[5]

(2a)	*gaṃṃaḷ*	'he got lice'	(2b)	*kammal*	'he completed'
	gaḅuḷ	'before'		*balad*	'country'
	ḅuṛaṃ	'he plaited'		*risam*	'he drew'

[1] From Sept 2008, Council for British Research in the Levant (Amman) & University of Salford.
[2] In Arabic transliteration, emphatics are denoted by a subscript dot.
[3] See e.g. Davis (1995), Shahin (2002), Younes (1993).
[4] Macrons denote long vowels.
[5] On the designation 'Muslim Baghdadi', see Blanc's (1964) *Communal Dialects in Baghdad*. This data from informants, but see also Woodhead & Beene's (1967) *Dictionary of Iraqi Arabic*.

- In (2a) the short low vowel *a* is backed, but in (2b) it is fronted / raised. Also, note the (predictable) *i~u* alternation in the past-tense triliteral verb.
- The consonants in (2a) are typically analysed as 'emphatic', although not lexically and there is no 'primary' emphatic causing the 'emphasis'.

1.3 The problem

- If 'emphasis' is defined as the spread of A (or [+pharyngeal], etc), then the examples in (1–2) above either cannot be 'emphasis', or we have to posit underlying non-coronal emphatics.
- Muslim Baghdadi Arabic (MB) – many examples such as (2a) in which 'back' quality not arbitrary, but mostly predictable.
- Moreover, MB shows that these 'secondary emphatics' do not spread their emphatic quality into non-emphatic coronal obstruents (so are unlike coronal obstruents in this respect) – this causes problems if we see 'emphasis' as A-spread:

(3)

taṛṛaḅ	'he got (sth) dirty / dusty'
mdaḅḅaṛa	'well organised (f.s.)'
yisṃaṛṛ	'it turns brown / gets tanned(m.s.)'

- Examples such as (1) above can be shown also to be generally predictable (with the possible exception of *l̩*, which I argue has a historical cause subsequently lexicalised).
- Too much regularity in the consonantal environment for this 'backing' to be due to underlying vowel contrasts.

2 Segments and elements

- Segmental representations are given in the appendix (contour structure, adapted from Nasukawa & Backley, 2005).
- As per Bellem (2007), the elements I assume are A, I, U, L, H, ?.
- Roughly, I is palatal, U is labial and A is pharyngeal.
- A does not characterise coronals – coronals and velars are said by Nasukawa & Backley (2005) to be differentiated by structure, thus having no resonance element. However, I claim that Arabic coronals are front consonants, thus characterised by the presence of I.
- Arabic consonantal system characterised overall by resonance qualities (cf. 'front'–'back' contrasts) and coronals characterised by presence of I (since they are 'front' consonants). This is what makes Arabic-style coronal emphatics 'marked' (in having two resonance qualities associated).

3 Emphatics in Muslim Baghdadi Arabic

3.1 Vowel system

- Short Vs in Arabic seen by the grammarians as *ħarakāt* (lit. 'movements', i.e. transitions) – notoriously unstable.
- Classical Arabic (CA) had *i a u* – this inventory varies across dialects (?mid-Vs / merging of high Vs).
- MB has *i a u* (although *i-u* merged in many contexts and actual realisation dependent on consonantal context).
- CA had 'long vowels' *ī ā ū* and diphthongs *ay aw*.
- MB has long *ī ē ā ō ū* (rarely *ay aw*).

3.2 Emphatic spread

3.2.1 Root-internal

- A majority of Arabic words based on tri-consonantal system ('radicals') – templatic morphology works on permutations of these consonants into different syllable templates (with some positions associated to other Cs and Vs):[6]

(4) K-T-B: semantic notion of 'writing'

KaTaBa	'he wrote'	*KuTiBa*	'it was written (m.s.)'
KaTaBat	'she wrote'	*KaTaBtu*	'I wrote'
yaKTuBu	'he writes'	*KaTTaBa*	'he made (s.o.) write'
istaKTaBat	'she dictated'	*taKāTaBū*	'they (m.) corresponded'
KiTāB(un)	'book' (indef. nom.)	*KāTiB(un)*	'writing / a writer' (indef. nom. m.s.)
maKTūB(un)	'written / a letter' (indef. nom. m.s.)	*maKTaB(un)*	'office / study' (indef. nom.)

- Leftward vs rightward spread of ('primary') coronal emphatics – stem-internally, MB (like most dialects) has C-C, Ṣ-Ṣ, Ṭ-C ... but C-Ṭ is rare (I found no examples in MB):

(5)

Ṣ-C		Ṭ-Ṭ	
ṣāmit-a	'silent (f.s.)'	*ṣiraṭ*	'he gulped (food)'
ṭumas	'he got bogged down'	*ṣiṭar*	'he chopped (meat)'
ṭumas-at	'she got bogged down'	*ṭawwaṭ*	'he hooted (the horn)'
ṭass-a	'hole (in the road)'	*ṣimīṭ*	'simit (like bagel)'
ṭaqis	'weather'	*ṣaṭal*	'bucket'

- Is this spread synchronic? Assumed so from orthography. If it were, this spread within the root probably at a root (consonantal) level.

3.2.2 Beyond the root

- Does the emphatic property spread into affixal consonants?
- The data in the literature vary on this point; to some extent this may be dialectal variation, but my data suggest that this spread is optional – it seems to be segmental and assimilatory (especially leftwards), suggestive of phonetic coarticulation.[7]

(6) Prefixes

sta-wṭan	[stawtʌn] ~ [stōṭʌn]	'he settled down'
t-waḍḍaf	[ˈtwʌðˠðʌf]	'he got a job'
ta-wṣiya	[tawsˠijɛ]	'(on) commission'
t-ḍaxxum	[təðˠʌχχʊm] ~ [ˀðˠðˠʌχχʊm]	'you (m.s.) inflate (sth)'
da-t-ḍaxxum	[dɛðˠðˠʌχχʊm]	'you're inflating (sth)'
yi-t-ṭāfaḥ	[jɪtˠtˠɑ:faħ]	'it's overflowing (m.s.)'
l-ṭifil	[ˀtˠɪfɪl]	'the child'
l-ṣurṣur	[ˀsˠʊrsˠʊr]	'the cockroach / cricket'

[6] The data are for Classical Arabic. Contemporary dialects differ from CA and each other in certain aspects.
[7] Some of the experimental literature supports this. See, e.g. Zawaydeh (1998) on gradient effects.

(7) Suffixes

ḍaxxam-t	[ðʌχχʌmɪt]	'I inflated (sth)'
γilaṭ-tu	[ʁɻlʌṭtu] ~ [ʁɻlʌṭt̪ʊ̪]	'you (pl.) made a mistake'
gambaṣ-ti	[gʌmbʌṣti] ~ [gʌmbʌṣtʷɪ][8]	'you (f.s.) squatted'

3.2.3 'Emphasis'

- There are cases analysed as 'emphasis' not covered by the data above. These seem to vary cross-dialectally. Sometimes apparent 'emphasis' perceptible in the consonants, but most often cued by vocalic contrasts.
- MB particularly interesting here because of the sheer number of clearly 'backed' words / syllables that don't fit the usual analyses.
- MB doesn't have such salient 'fronting' as some other dialects, so that in other dialects the effect of 'fronting' may be the more salient contrast. Generative literature – no investigations of 'emphasis' in context of 'fronting', although the grammarians recognised 'fronting' effects (termed *imāla* 'inclination [of *a* towards *i*]').[9]
- The following section looks at the behaviour of other consonants in MB.

4 Other 'back' contexts in Muslim Baghdadi Arabic

- It is instructive to look at other consonants in MB, as this demonstrates the problem with analysis of 'emphasis'.
- One typical place to look for variation is in suffixes. In MB the 'feminine' (noun) ending is *-a(t)*; variation is commonly seen in the realisation of the low vowel – in a front context (under *imāla*) it is [ɛ] (when unstressed it is not very stable, often sounds closer to schwa [ə]); after a pharyngeal it is close to [a]; in a 'back' / emphatic context it is closest to [ɐ] (i.e. a reduced 'non-front' *a*):

(8) MB 'feminine' suffix *-a*

a. 'front' context [-ɛ]

karrad-a	Karrada (district in Baghdad)
ṣāmit-a	'silent (f.s.)'
ṭass-a	'hole (in the road)'
gamz-a	'(a) jump'

b. pharyngeal context [-a]

rābiʕ-a	'fourth (f.s.)'
maṭrūḥ-a	'spread out (f.s.)'

c. emphatic context [-ɐ]

baṭṭ-a	'(a) duck'
mšāṭ-a	'(hair) combs'
gaṣgūṣ-a	'a piece / scrap'
mugrāḍ-a	'nail clippers'

[8] For both these examples, my informant produced the suffixes as emphatic only in citation form and preferred my realisations non-emphatic.

[9] Particularly Sībawayh (d. late 8th C.). See al-Nassir (1993: *passim*).

- Stem-final consonants can also be tested with the objectival suffixes: masculine singular *-a* and feminine singular (also inanimate plural) *-ha*:

(9)	*kitab-a*	'he wrote it (m.s.)'	*kitab-ha*	'he wrote it/them (f.s./pl.)'
	buṣaṭ-a	'he hit him'	*buṣaṭ-ha*	'he hit her'

- These are the same as the possessive suffixes: masculine singular *-a* and feminine singular (also inanimate plural) *-ha*:

(10)	*bēt-a*	'his house'	*bēt-ha*	'her house'

4.1 Labials

- Stem-finally, labials seem to 'back' when there is no overt 'front' context adjacent:

(11)	*guṣb-a*	[gʊ̟ṣbɐ]	'straw'
	ḏ̪arb-a	[ð̟ʌrbɐ]	'(a) hit, blow'
	ǰarb-a	[dʒarbɐ] ~ [dʒɛrbɐ]	'mangy (f.s.)'
	galb-a	[gʌɬbɐ]	'his heart'
	labw-a	[lɛbwɐ]	'lionness'
	rugb-a	[rʊgbɐ]	'neck'
	gahw-a	[gahwɐ] ~ [gɑhwɐ]	'coffee'
	karam-a	[karʌmɐ]	'his generosity'
	kahrab-a	[kahrʌbɐ]	'it electrocuted him'
	waḏ̪ḏ̪af-a	[wʌð̟ð̟ʌfɐ]	'he employed him'
	ṭabb-a	[tˠʌbbɐ]	'he entered it'
	maħbūb-a	[maħbu:bɐ]	'adorable (f.s.)'

- Where there is an adjacent 'front' context, however, this 'palatalises' stem-final labials:

(12)	*ṣudf-a*	[sˠʊdfɛ]	'coincidence'
	kitab-a	[kɪtɛbɛ]	'he wrote it'
	ṭabīb-a	[tˠʌbi:bɛ]	'doctor (f.s.)'

- Non-final labials in a front environment:

(13)	*bēḏ̪*	[biʌð̟] ~ [biɐð̟]	'eggs'
	ywaswas	[jwɛswɛs]	'he whispers'
	ballal	[bɛllɛl]	'he / it got wet'

- Medial labials in a 'conflicting' environment:

	gabban	[gʌbban]	'he weighed'
	gammaz	[gʌmmaz]	'he leapt around'
	gamz-a	[gʌmzɛ]	'(a) jump'

- So, in combination with *r* (where no *i* is present) or *g* (where no *i* or front consonant is present) labials seem to be back (in that they yield [ʌ]), but they seem to become fronted ('palatalised') where a front consonant or vowel is present
- Labials are U consonants which may be effectively 'palatalised' (i.e. they can participate in front domains), but they are not 'emphatic'

- This means that labials are in a sense 'back' (when not 'palatalised') but they do not become emphatic

4.2 Velars

- Word-finally, *k* is 'front', and *g* is most often 'front' (exceptions discussed presently):

(14)

ragg-a	[rʌggɛ]	'turtle'
šagg-a	[ʃɛggɛ]	'he tore it up'
gamrag-ha	[gʌmraghɛ]	'he customs-inspected it (f.s.)'
dabk-a	[dɛbkɛ]	'dabka (dance)'
ṭagg-a	[tˤʌggɛ]	'a bang / burst'

- In other positions, *g* and *k* behave somewhat differently – *g* participates in 'back' domains, where *k* doesn't:

(15)

a.

gabban	[gʌbban]	'he weighed'
gammaz	[gʌmmaz]	'he leapt around'
gamz-a	[gʌmzɛ]	'(a) jump'
gammaḷ	[gʌmmʌl]	'he got lice'
gahw-a	[gahwɐ] ~ [gɑhwɐ]	'coffee'

b.

kammal	[kɛmmɛl]	'he completed'
karrad-a	[kɛrrɑ:dɛ] ~ [karrɑ:dɛ]	Karrada (district)
karrah	[kɛrrah] ~ [karrah]	'he made hate'
kabbar	[kɛbbʌr] ~ [kabbʌr]	'he made bigger / enlarged'

- Where there is a 'front' context, however, then (as word-finally) *g* participates:

(16)

gadd	[gɛdd]	'equal to'
gaššar	[gɛʃʃar] ~ [gɛʃʃʌr]	'he peeled'
gēmar	[giɛmʌr]	'(buffalo) cream'

- From this, it seems that *g* has no inherent resonance (lexically unspecified), whereas *k* seems inherently I-specified (phonologically palatalised).
- Consonants seem to have a noticeably leftward effect on resonance.
- However, there is some rightward effect – with stem-final *g* a U-domain is created where a labial precedes (including long Vs),[10] and this seems to prevent the stem-final I effect:

(17)

(*bēḍa*) *maṭgūg-a*	[mʌtgu:gɐ]	'broken (egg) (f.s.)'
bōg-a	[bo:gɐ]	'(a) burglary'
bagg-a	[bʌggɐ]	'(a) mosquito'

- This is similar to the labial case.
- With the velar, the prevention of the stem-final I effect occurs with labials but not emphatics (*ṭagg-a* '(a) bang') (as if *g* takes on the U quality, but not the A quality).

[10] Lexically, I argue that Arabic 'long vowels' are in fact VCV.

4.3 Liquids

- The liquids *r* and *l* behave slightly differently. *r* is 'back' unless adjacent to *i / ī*, when it is 'de-backed':

(18)

marīḍ	[mɛri:ð]	'ill (m.s.)'
farid	[fɛri:d]	'Farid (name) (m.s.)'
fard-a	[fʌrdɛ]	'one of a pair'
manbar-a	[mɛnbʌrɐ]	'his pulpit'
idār-a	[ɪdɑ:rɐ]	'administrative office'
dīr-a	[di:rɛ]	'district'
dōr-a	[do:rɐ]	'his turn / role'
gadar	[gɛdʌr] ~ [gɛdar]	'measure(ment)'

- Note: in *farda*, the adjacency of a front consonant has not fronted the *r*.
- With *ē*, where a preceding labial or velar (likewise a non-emphatic coronal) palatalises (which is a noticeable feature of this dialect), *r* (like emphatic coronals and gutturals) does not palatalise: *rēl* [rɛ:l].
- Thus, *r* is 'back' and similar to emphatic (I claim that it has an inherent A), although unlike the emphatic coronal obstruents it (1) is easily de-emphaticised, and (2) doesn't spread its emphatic quality into other coronal obstruents. Thus, true 'emphatic-ness' is a property of the obstruent system, while these resonance patterns affect the whole of the sound system.
- The liquid *l* is most often front (inherent I), although it is susceptible to backing in an otherwise 'back' context. There are also cases where it seems lexically 'back' (this has a historical basis). Cf. *labb-a* 'a kick', *lablab* 'he extracted the core', *rabb-a* 'his lord', *rabrab* 'he kicked up a hell of a fuss'.

5 Back to *imāla*: resonance domains in Arabic

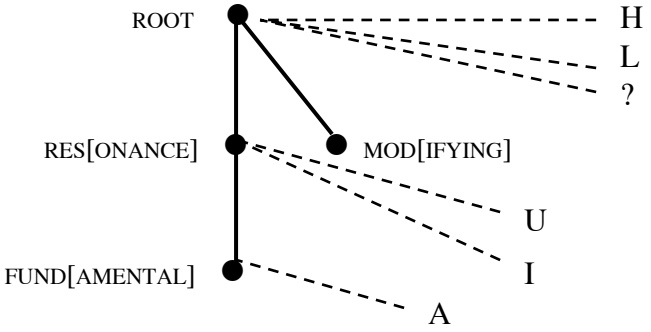
- Data above appears complex; however, if we stop looking at one active process of 'emphasis' we can see better how the entire sound system patterns (at least, in terms of resonance).
- As above, the grammarians talked of 'fronting' effects on the low vowel (*imāla*).
- Proposal for coronals in Arabic being characterised by element I fit neatly with this. That is, I-spread (*imāla*) is an active synchronic process affecting non-A (non-emphatic) domains.
- Emphatics *also* have an A element (which may spread from obstruents – sonorants are 'weaker').
- Some dialects (such as MB) also have U domains in which I is not active. These 'domains' are where elements seem active at a suprasegmental level (the CV level).
- Compare *mabrūma / mablūl* – the latter shows the active I-spread (or rather, 'I domain').
- This raises the question of what exactly we mean by 'emphatic' and 'emphasis' – I argue that we should look at the overall sound system (in this case, at the resonance patterning).

Appendix

Segmental representations

- These segmental representations (with contour structure) roughly adapted from Nasukawa & Backley (2005) (henceforth N&B), for reasons argued in Bellem (2007).
- As per Bellem (2007), the elements adopted are A, I, U, L, H, ?.
- This is similar to Revised Element Theory, whereby L is both 'voicing / low tone' and 'nasality' and H is both 'aspiration / high tone' and 'frication' – both thus have the dual function of 'manner' / 'laryngeal setting' (see (5) below) – ? similarly has dual function ('glottalisation' and 'stop-ness').
- Note that A does not characterise coronals (for N&B, coronals and velars are differentiated by structure):

(19) Consonantal representations



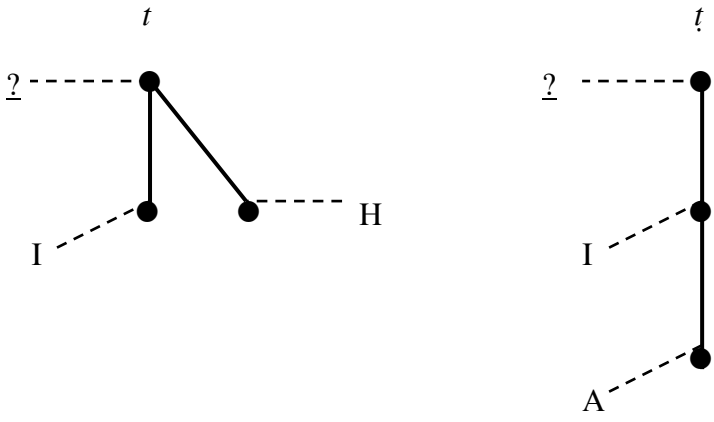
Notes to representation in (19):

(i) *For N&B, coronals have a structure terminating in a bare RES NODE and velars in a bare FUND node*

(ii) *The ROOT node is interpreted as manner (thus maximally consonantal)*

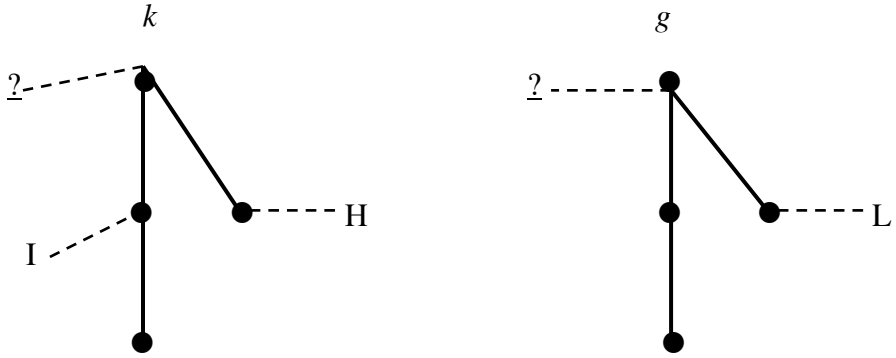
(iii) *Elements have dual function, depending on the node to which they attach – H, L, ? are interpreted as 'manner' on the ROOT node, but as 'modifying' (roughly, laryngeal setting) on the MOD node.*

(20) Coronal obstruents in Arabic



(21) Arabic liquids

(22) Arabic velars



References

Bellem, Alex (2007). 'Towards a Comparative Typology of Emphatics: Across Semitic and into Arabic Dialect Phonology'. PhD thesis, SOAS, University of London.

Blanc, Haim (1964). *Communal Dialects in Baghdad*. Cambridge, MA: Harvard University Press.

Davis, Stuart (1995). 'Emphasis Spread in Arabic and Grounded Phonology'. *Linguistic Inquiry* 26: 465–98.

al-Nassir, A.A. (1993). *Sibawayh the Phonologist: A Critical Study of the Phonetic and Phonological Theory of Sibawayh as Presented in his Treatise Al-Kitab*. London / New York: Kegan Paul International.

Nasukawa, Kuniya & Phillip Backley (2005). 'Dependency Relations in Element Theory'. *Leiden Papers in Linguistics*. 2.4: 77–93 (special issue: Proceedings of the Government Phonology Workshop, eds N. Kula & J. van de Weijer).

Shahin, Kimary N. (2002). *Postvelar Harmony*. Amsterdam: John Benjamins.

Woodhead, D.R. & Wayne Beene, eds (1967). *A Dictionary of Iraqi Arabic*. Washington, DC: Georgetown University Press.

Younes, Munther (1993). 'Emphasis Spread in Two Arabic Dialects'. *Perspectives on Arabic Linguistics V.* (eds) M. Eid & C. Holes. Amsterdam: John Benjamins.

Zawaydeh, Bushrah A. (1998). 'Gradient Uvularization Spread in Ammani-Jordanian Arabic'. *Perspectives on Arabic Linguistics XI.* (eds) A. Benmamoun, M. Eid & N. Haeri. Amsterdam: John Benjamins.