University of Massachusetts Amherst

Prosodic Augmentation of the Moroccan Arabic Broken Plural

Ali Nirheche & Michael Becker

RFP: July 1, 2025

Broken plural patterns in Moroccan Arabic

- Two kinds of plurals in Moroccan: “sound” = suffixal and **“broken” = templatic**
- The Nirheche (2025) corpus: 891 plurals (45% broken), based on the Darija Open Dataset (Outchakoucht & Es-Samaali 2021).
- Only 3 broken plural patterns are productive in Moroccan, cover 95% of the data.

Pattern	**%**	**σσ plurals**	**σσσ plurals**
C.CaC…	60%	[b.nat] ‘women’	[m.sa.kən] ‘paupers’, [l.ja.li] ‘nights
C.Cu.C…	27%	[q.run] ‘centuries’	[k.tu.ba] ‘books’
CX.Can	8%	[ki.san] ‘sacks’	—
Others	5%	kuħəl ‘blue.pl’, …	—

Highlights

- Moroccan Arabic **σσ** broken plurals are augmented to **σσσ**:
  - Stage 1, inherited: nouns take **σσ** broken plurals, ALIGNR-ROOT
  - Stage 2, intermediate: **feminines** start taking **σσσ,** driven by MAX
  - Stage 3, present: **masculines** start taking **σσσ**, driven by NONFINALITY
  - Stage 4, possible future: **all nouns** will take **σσσ** plurals
- All broken plurals start with an iamb = stress on the second syllable
  - final stress in (**σ**ˈ**σ**) violates NONFINALITY
  - non-final stress in (**σ**ˈ**σ**)**σ** satisfies NONFINALITY
- Changes in constraint weight cause diachronic change and govern synchronic variability
- Sources of data: our corpus, Harrell et al. (1966)'s dictionary, a survey with 42 speakers

Stage 1 to stage 2: plural C.CəC becomes C.Ca.Ci

- C.Ca.Ci plurals are inherited (Classical CaCa:Ci:, e.g. [la.ja:.li:] ‘night’)
- In our corpus: only 5 out of 40 (12%) C.Ca.Ci plurals come from Classical Arabic
- **Feminine** pl CCəC (Classical CiCaC, CuCaC) → C.Ca.Ci

	SG	Classical	Stage 1	Stage 2	
stuck in stage 1	qəl.la	qi.lal	q.ləl	—	‘clay jug’
variable	rəz.ma	ri.zam	r.zəm	r.za.mi	‘bundle’
stage 2 only	ħək.ma	ħi.kam	—	ħ.ka.mi	‘wisdom’
	rək.ba	ru.kab	—	r.ka.bi	‘knee’

Stage 1 to stage 2: plural C.CəC becomes C.Ca.Ci

- Stage 1: C.CəC plurals optimize AlignR-Root, but has final stress
- Stage 2: **Feminine** plural C.CəC → C.Ca.Ci
- Max gains weight and prevents the deletion of the feminine suffix [a]

Stage 1

r	ə	z		m	a
r		z	ə	m	

Max	NonFinality	AlignR-Root
1	1	0

[r.zəm] has [a]-deletion, final stress

Stage 2

r	ə	z		m	a
r		z	a	m	i

Max	NonFinality	AlignR-Root
0	0	1

[r.za.mi] has no deletion,no final stress

Stage 2 to stage 3: C.CuC, CCaC → C.Cu.Ca

- **Masculines**: C.CuC, C.CaC are changing to C.Cu.Ca to remove NonFinality violation
- C.Cu.Ca is a recent innovation of North African Arabic (Heath 1987; Ratcliffe 2002). No corresponding pattern in Classical Arabic (*CVCu:Ca:)

SG	stage 2	stage 3	PL	
qəl.b	C.CuC →	C.Cu.Ca	q.lub ~ q.lu.ba	'heart'
b.ħər			b.ħur ~ b.ħu.ra	'sea'
r.bəʕ	C.CaC →		rˤ.baʕ ~ rˤ.bu.ʕa	'quarter'
ʕ.dˤəm			ʕ.dˤam ~ ʕ.dˤu.ma	'bone'

Stage 2 to stage 3: C.CuC, CCaC → C.Cu.Ca

- **Masculines**: NonFinality gains more weight, overpowers AlignR-Root

Stage 2

q ə l b
q l u b

Max	NonFinality	AlignR-Root
0	1	0

[q.lub] has final stress

Stage 3

q ə l b
q l u b a

Max	NonFinality	AlignR-Root
0	0	1

[q.lu.ba] has epenthesis, non-final stress

Diachronic analysis: stage 1 (inherited)

- The MaxEnt grammar prefers **σσ** plurals for both **masculine** and **feminine**.
- Only AlignR-root is active at this stage, preventing final vowels from surfacing.

		Max $w=0$	NonFinality $w=0$	AlignR-Rt $w=10$	$\mathscr{H}$	p
/qəlb/ + /PL/	**q.lub**		−1		0	**.99**
	q.lu.ba			−1	−10	.01
/rəkba/ + /PL/	**r.kəb**	−1	−1		0	**.99**
	r.ka.bi			−1	−10	.01

Diachronic analysis: stage 2 (intermediate)

- Raising MAX (0 ⟶ 10) rewards feminine C.CV.CV that keep the final vowel.
  The grammar prefers both **σσ** and **σσσ** plurals for **feminines**.
- **Masculines**: no change

		↑MAX w=10	NONFINALITY w=0	ALIGNR-RT w=10	$\mathscr{H}$	p
/qəlb/ + /PL/	**q.lub**		−1		0	**.99**
	q.lu.ba			−1	−10	.01
/rəkba/ + /PL/	**r.kəb**	−1	−1		−10	**.5**
	r.ka.bi			−1	−10	**.5**

Diachronic analysis: stage 3 (today)

- Raising NonFinality (0 ⟶ 10): tie between **σσ** and **σσσ** for **masculines**.
  The grammar prefers both **σσ** and **σσσ** plurals for masculines
- **Feminines**: strongly prefer σσσ.

		Max w=10	↑NonFinality w=10	AlignR-Rt w=10	$\mathscr{H}$	p
/qəlb/ + /pl/	**q.lub**		−1		−10	**.5**
	q.lu.ba			−1	−10	**.5**
/rəkba/ + /pl/	r.kəb	−1	−1		−20	.01
	r.ka.bi			−1	−10	**.99**

Diachronic analysis: stage 4 (a possible future)

- Only **σσσ** broken plurals in Moroccan
- Raising NonFinality (10 → 20) rewards (**σˈσ**)**σ** without final stress

		Max w=10	↑NonFinality w=20	AlignR-Rt w=10	$\mathscr{H}$	p
/qəlb/ + /pl/	q.lub		−1		−20	.01
	q.lu.ba			−1	−10	**.99**
/rəkba/ + /pl/	r.kəb	−1	−1		−30	.01
	r.ka.bi			−1	−10	**.99**

Diachronic analysis: further detail

- Not all patterns/items are changing at the same rate:
  - Stage 2 left very few stragglers (~3 items stuck in C.CəC)
  - Stage 3: changes are happening at different rates:
    - C.CuC → C.Cu.Ca is strong: 22 out of 54 (41%) items
    - C.CaC → C.Cu.Ca is weaker: 4 out of 41 (10%) items
  - Stage 4 (the future) is already here for many words!
    - C.CuC → C.Cu.Ca is complete for 22 nouns
- More detail in our complete analysis (not presented here):
  - Includes **vowel quality**, e.g. feminine C.Ca.Ci vs. masculine C.Cu.Ca
  - It accounts for **exceptional items** using indexed constraints (Pater 2000, 2007, 2010)

Case study: CCuC ➜ CCuCa

- We studied CCuC(a) plurals to investigate this ongoing change: CCuC ➜ CCuCa
- We extracted all nouns that take CCuC(a) plural from the corpus: 86 nouns.

	status of [a]	example		*n*
a.	No [a]	ʒ.durˤ	'roots'	42
b.	Optional [a]	w.ʒuh ~ w.ʒu.ha	'faces'	22
c.	Obligatory [a]	n.mu.ra	'tigers'	22

Recent expansion of C.Cu.Ca

- A comparison with Harrell et al.'s (1966) dictionary: C.Cu.Ca is increasing

	contemporary corpus		
	No [a]	Optional	With [a]
Harrell et al.			
No [a]	26	10	—
Optional	—	12	6
With [a]	—	—	9

Survey

- We conducted a study to generate a more nuanced understanding of the distribution of final [a] in C.CuC(a) plurals
- **Participants:** 42 native speakers of Moroccan Arabic

Survey: materials

- **Materials:**
  - 18 nouns with CCuC(a) plurals selected from the corpus: 4 items with no [a], 10 with optional [a], and 4 with obligatory [a]
  - Each noun was presented within a frame sentence in Arabic script with emojis, followed by a question asking participants to choose which plural (C.CuC or C.Cu.Ca) sounded better
- **Procedure:**
  - The experiment was distributed online using Experigen (Becker & Levine 2015)

Survey

The king has a big **qsʕər**

The king has many _____

Which plural sounds better to you?

qsʕur | qsʕura

Old people say qsʕur, not qsʕura

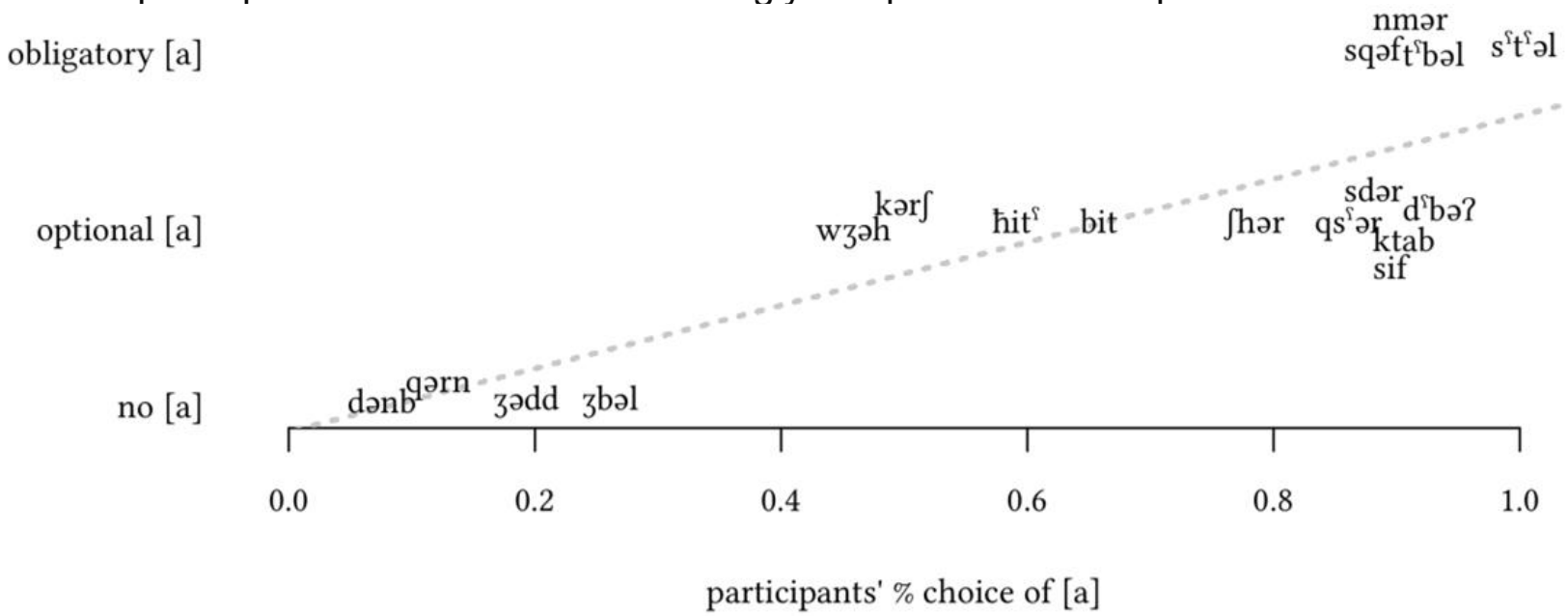
true | false

Women say qsʕur, not qsʕura

true | false

Survey: results

- The selection of the final [a] is gradient across the 18 items
- Participants preferred C.Cu.Ca more strongly compared to the corpus



Support from prior work on Moroccan Arabic

- **Boudlal (2001):** diminutives also undergo augmentation to avoid **(σˈσ)**
- Augmentation strategies: final [a] or internal schwa

				sg		diminutive	
Fem	C.CiC	→	C.Ci.Ca	bən.t	→	b.ni.ta (*b.nit)	'girl'
Masc	C.CiC	→	C.Ci.jəC	kəl.b	→	k.li.jəb (*k.lib)	'dog'

- **Ratcliffe (2002):** Moroccan Arabic broken plurals prefer **σσσ**
- Our analysis: preference for **σσσ** modeled by re-weighting of NonFinality and Max.

Conclusion

- Broken plurals in Moroccan Arabic are moving from **σσ** to **σσσ**;
    - Feminines began the change (Stage 2),
    - Masculines are currently changing (Stage 3),
    - Possible future: all plurals might become σσσ (Stage 4).
- We analyzed the diachronic increase in augmentation by increasing the weight of MAX first, then the weight of NONFINALITY.
- Evidence comes from three sources: 1960s dictionary, corpus (891 plurals), and a survey with 42 speakers.
- Future work: test the predictions of our grammar with nonce words. We predict a preference for **σσσ** plurals, more strongly for feminines

Thank You

University of Massachusetts Amherst

Moroccan Arabic plurals

- Two kinds of plurals in Moroccan: “sound” = suffixal (a) and “broken” = templatic (b)

		SG	PL	type	*n*		
a.	sound	l.san	l.sa.n-at	-at	395	81%	‘tongue’
		jəd.d	jəd.d-in	-in	63	13%	‘hand’
		ʃi.fur	ʃi.fu.r-a	-a	28	6%	‘driver’
		total sound:			486	100%	
b.	broken	kəl.b	k.lab	a-iamb	266	66%	‘dog’
		qər.n	q.run	u-iamb	86	21%	‘century’
				others	53	13%	
		total broken:			405	100%	

Corpus study

- The corpus: Nirheche (2025), based on the Darija Open Dataset (Outchakoucht & Es-Samaali 2021).
- Contains 891 plurals with their corresponding singulars in IPA, of which 405 (45%) are broken.
- The broken plural patterns that were or are involved in diachronic changes:

template	n	SG	PL	
C.CuC(a)	86	qəl.b	q.lub(a)	‘heart’
C.CaC	70	kəl.b	k.lab	‘book’
C.Ca.Ci	40	rək.ba	r.ka.bi	‘knee’
C.CəC	2	tˤa.sˤa	tˤ.jəsˤ	‘container’

Today's focus: plurals of small singular nouns

- We focus on **small** singular nouns, i.e., those with the singular pattern **CXC(a)**:

Small masculine	**Small feminine**
CəC.C, C.CəC	CəC.Ca
CiC	CiCa
CuC	CuCa
CaC	CaCa

Diachronic changes: stage 1

- **Masculine** nouns (CCC/CVC) have disyllabic plurals: CCaC, CCuC

	CA pl.pattern	MA pl.pattern	example	
a.	ʔaCCa:C	CCaC	wəld → wlad	'boys'
	CiCa:C	CCaC	kəlb → klab	'dogs'
b.	CuCu:C	CCuC	qəlb → qlub	'hearts'
	ʔaCCuC	CCuC	wʒəh → wʒuh	'faces'

Diachronic changes: stage 1

- **Feminine** nouns (CCCa/CVCa) have disyllabic plurals: CCaC, CCəC

	CA pl.pattern	MA pl.pattern	example	
a.	ʔaCCa:C	CCaC	mətla → mtal	‘proverbs’
	CiCa:C	CCaC	reqba → rqab	‘dogs’
b.	CiCaC	CCəC	rəzma → rzəm	‘bundle’
	CuCaC	CCəC	nəqba → nqəb	‘hole’

Diachronic changes: stage 2 (intermediate)

- **Masculine** nouns (CCC/CVC) show no change.
- **Feminine** nouns (CCCa/CVCa) with expected CCəC have CCaCi

	CA pl.pattern	MA pl.pattern	example	
a.	CiCaC	CCaCi	ħəkma → ħkami (*ħkəm)	'wisdom.PL'
			firaq → fərqa (*frəq)	'groups'
b.	CuCaC	CCaCi	rəkba → rkabi (*rkəb)	'knees'
			ħəfra → ħfari (*ħfər)	'holes'

Diachronic changes: stage 2 (intermediate)

- C.Ca.Ci plurals have a corresponding CA pattern (CaCa:Ci:) but it's not productive in CA
- C.Ca.Ci extended beyond their CA origins to cover nouns with the productive CVCaC pattern.
- Only 5 out of 40 (12%) C.Ca.Ci plurals in our corpus have a CA source.

	singular	Moroccan plural	MSA plural	
a.	dər.ri	d.ra.ri	ða.ra:.ri:	'boy'
	li.la	l.ja.li	la.ja:.li:	'night'
b.	rək.ba	r.ka.bi	ru.kab	'knee'
	fər.qa	f.ra.qi	fi.raq	'team'

Diachronic changes: stage 3 (today)

- Masculine nouns (CCC/CVC) are currently adopting CCuCa: we observe within-word variation (CCuC~CCuCa) in today's grammar.
- CCuCa is a recent innovation of Moroccan Arabic (Heath 1987; Ratcliffe 2002). There is no corresponding pattern in Classical Arabic (*CVCuuCaa).

	change			SG		PL	
a.	C.CuC	→	C.Cu.Ca	qəl.b	→	q.lub ~ q.lu.ba	'hearts'
				k.tab	→	k.tub ~ k.tu.ba	'books'
b.	C.CaC	→	C.Cu.Ca	r.bəʕ	→	rˤ.baʕ ~ rˤ.bu.ʕa	'quarters'
				ʕ.dˤəm	→	ʕ.dˤam ~ ʕ.dˤu.ma	'bones'

Diachronic changes: stage 3 (today)

- Most feminine nouns (CCCa/CVCa) already fully adopted CCaCi
- Few remaining nouns are currently changing.

	change			SG		PL	
a.	C.CuC	→	C.Ca.Ci	həd^{ʕ}.r^{ʕ}a	→	h.d^{ʕ}ur^{ʕ} ~ h.d^{ʕ}a.r^{ʕ}i	'talks'
				kəs.ra	→	k.sur ~ k.sa.ri	'piece of bread'
b.	C.CaC	→	C.Ca.Ci	rəq.ba	→	r.qab ~ r.qa.bi	'nape'
				χər.s^{ʕ}a	→	χ.ras^{ʕ} ~ χ.ra.s^{ʕ}i	'ear ring'

C.Cu.Ca encroaching on C.CaC

- C.CaC → C.Cu.Ca, driven by NonFinality, even at the cost of Ident(high) and Dep.

singular	Harrell et al.	contemporary plural	
r.bəʕ	r.baʕ ~ r.bu.ʕa	r.bu.ʕa	‘quarter’
dˤ.bəʕ	dˤ.baʕ ~ dˤ.bu.ʕa	dˤ.bu.ʕa	‘hyena’
ʕ.dˤəm	ʕ.dˤam ~ ʕ.dˤu.ma	ʕ.dˤu.ma	‘bone’
tˤərˤ.f	tˤ.rˤaf	tˤ.rˤu.fa	‘fraction’
ʒbəl	ʒ.bal	ʒ.bal ~ ʒ.bu.la	‘mountain’

- Changes are unidirectional, always towards more [a], suggesting an ongoing diachronic change.

Analysis: the quality of the epenthesized vowel

- Epenthetic [a], no schwa in open syllable, OCP(high) eliminates [i, u]

/noun + u_{PL}/		*ə]$_\sigma$ $w = 5$	OCP(high) $w = 5$	$\mathscr{H}$	p
/kər.ʃ/	(k.ru)ʃa			0	1
	(k.ru)ʃi		−1	−5	0
	(k.ru)ʃə	−1		−5	0

Analysis: MaxEnt with indexed constraints

- We use MaxEnt (Goldwater & Johnson 2003) with lexically-indexed constraints (Pater 2000, 2007, 2010)
- Optionality of final [a] as a competition between NonFinality and Dep

/noun + u_{PL}/		NonFinality $w = 0$	Dep $w = 0$	$\mathscr{H}$	p
/kər.ʃ/	(k.ruʃ)	−1		0	.50
	(k.ru)ʃa		−1	0	.50

Analysis: simulation

- Software: Shiny app (Nirheche 2024), that is based on Harmonic Grammar in R (HGR, Staubs 2011) to learn the weights of the constraints.
  - **Training data:** the 67 words from the corpus
  - **Constraints:** NonFinality, Dep and indexed versions of each for every lexical item
- Python script to generate candidates and indexed constraints.

Analysis: results

- For words with optional [a], the model assigned a small weight to the indexed DEP constraint.

		NONFIN $w = 16$	NONFIN$_{dərb}$ $w = 0$	DEP $w = 14.9$	DEP$_{dərb}$ $w = 1.1$	$\mathscr{H}$	p
/dərb/ + u_{PL}	(d.rub)	−1	−1			−16	.50
	(d.ru).ba			−1	−1	−16	.50

Analysis: results

- For words with obligatory [a], the indexed NonFinality constraint was given enough weight to overcome Dep

		NonFin $w = 16$	NonFin$_{nmər}$ $w = 6.9$	Dep $w = 14.9$	Dep$_{nmər}$ $w = 0$	$\mathscr{H}$	p
/nmər/ + u$_{PL}$	(n.mur)	−1	−1			−22.9	.01
	(n.mu).ra			−1	−1	−14.9	.99

Analysis: results

- For words with prohibited [a], a higher weight was assigned to the indexed DEP constraint

		NONFIN $w = 16$	NONFIN$_{qərn}$ $w = 0$	DEP $w = 14.9$	DEP$_{qərn}$ $w = 9$	$\mathscr{H}$	p
/qərn/ + u_{PL}	(q.run)	−1	−1			−16	.99
	(q.ru).na			−1	−1	−23.9	.01