Arabic Linguistics Society
رابطة اللسانيات العربية

Permeability and Stratification in the Moroccan Arabic Diglossic Lexicon

Ali Idrissi and Ali Nirheche

Qatar University

UMass

INDIANA UNIVERSITY

ASAL 39

University of Indiana – Mach 27-29, 2026

جامعة قطر
QATAR UNIVERSITY

University of Massachusetts Amherst

Aim

- This talk examines the behavior of the definite article *l-* in Moroccan Arabic (MA).
- We focus on a puzzling pattern involving the voiced alveopalatal [ž] (an affricate in certain dialects).
- We argue that this pattern provides strong evidence for a (MA) diglossic lexicon that is:

↳ ***stratified***

and/but

↳ **permeable**

The data..

l-assimilation in Arabic

- In Standard Arabic (StA), assimilation of *l-* applies before coronals:

UF	**SF**		**UF**	**SF**	
/l-tamr/	[ttamr]	‘dates’	/l-baab/	[lbaab]	‘door’
/l-dars/	[ddars]	‘lesson’	/l-kitaab/	[lkitaab]	‘book’
/l-samak/	[ssamak]	‘fish’	/l-maaʔ/	[lmaaʔ]	‘water’
/l-šams/	[ššams]	‘sun’	/l-fiil/	[lfiil]	‘elephant’
/l-ṣabr/	[ṣṣabr]	‘patience’	/l-qamar/	[lqamar]	‘moon’
/l-naas/	[nnaas]	‘people’	/l-ḥaliib/	[lḥaliib]	‘milk’
/l-rajul/	[rrajul]	‘man’	/l-ʕamal/	[lʕamal]	‘work’
/l-layl/	[llayl]	‘night’	/l-hawaaʔ/	[lhawaaʔ]	‘air’

l-assimilation in Arabic

- However, it fails to do so before [ž] (and its affricate variant)

UF	**SF**		
/l-žaaʔiza/	[lžaaʔiza]	***[žžaaʔiza]**	‘prize’
/l-žiha/	[lžiha]	***[žžiha]**	‘direction’
/l-žuhuud/	[lžuhuud]	***[žžuhuud]**	‘efforts’
/l-žuuʕ/	[lžuuʕ]	***[žžuuʕ]**	‘hunger’
/l-žawaab/	[lžawaab]	***[žžawaab]**	‘answer’
/l-žamal/	[lžamal]	***[žžamal]**	‘camel’
/l-žadiid/	[lžadiid]	***[žžadiid]**	‘new’
/l-žabal/	[lžabal]	***[žžabal]**	‘mountain’
/l-žubn/	[lžubn]	***[žžubn]**	‘cheese’
/l-židaar/	[lžidaar]	***[žžidaar]**	‘wall’
/l-žuθθa/	[lžuθθa]	***[žžuθθa]**	‘corpse’

The puzzle…

Variability

- While in StA, [ž] never triggers assimilation, it behaves differently in MA:
- It triggers it in some cases but fails to do so in other cases.

Standard Arabic			**Moroccan Arabic**			
/l-žamal/	[lžamal]	*[žžamal]	/l-žməl/	[žžməl]	*[lžməl]	'camel'
/l-žarab/	[lžarab]	*[žžarab]	/l-žərba/	[žžərba]	*[lžərba]	'scabies'
/l-žanuub/	[lžanuub]	*[žžanuub]	/l-žanuub/	*[žžanub]	[lžanub]	'south'
/l-žism/	[lžism]	*[žžism]	/l-žism/	*[žžism]	[lžism]	'body'

This variability is the puzzle

Previous accounts

Two main accounts

1) one treats non-assimilating forms as borrowings from Standard Arabic
2) the other treats the variation as phonologically conditioned.

The ‘borrowing’ account

- … by Freeman (2016), Harrell (1962) and Heath (1987)
- The key claim:
  - The exceptional forms are **borrowings** from StA
  - As such, they are **lexically specified** as exceptions to ***l-ž*** *assimilation*

Standard Arabic			**Moroccan Arabic**			
/l-žamal/	[lžamal]	*[žžamal]	/l-žməl/	[žžməl]	*[lžməl]	‘camel’
/l-žarab/	[lžarab]	*[žžarab]	/l-žərba/	[žžərba]	*[lžərba]	‘scabies’
/l-žanuub/	[lžanuub]	*[žžanuub]	/l-žanuub/	***[žžanub]**	**[lžanub]**	‘south’
/l-žism/	[lžism]	*[žžism]	/l-žism/	***[žžism]**	**[lžism]**	‘body’

[-coronal assimilation]

‘Phonological conditioning’ account

- Nirheche (2025) observed that the behavior of [ž] in MA is **gradient** (not categorical):
- ***l-ž*** *assimilation* occurs in:
  - 96% of cases when [ž] precedes a consonant
  - 84% of cases when [ž] is followed by a schwa
  - 37% of the cases when before a full vowel

/l-žməl/	[žžməl]	*[lžməl]	‘camel’
/l-žərba/	[žžərba]	*[lžərba]	‘scabies’
/l-žib/	[žžib]	*[lžib]	‘pocket’’

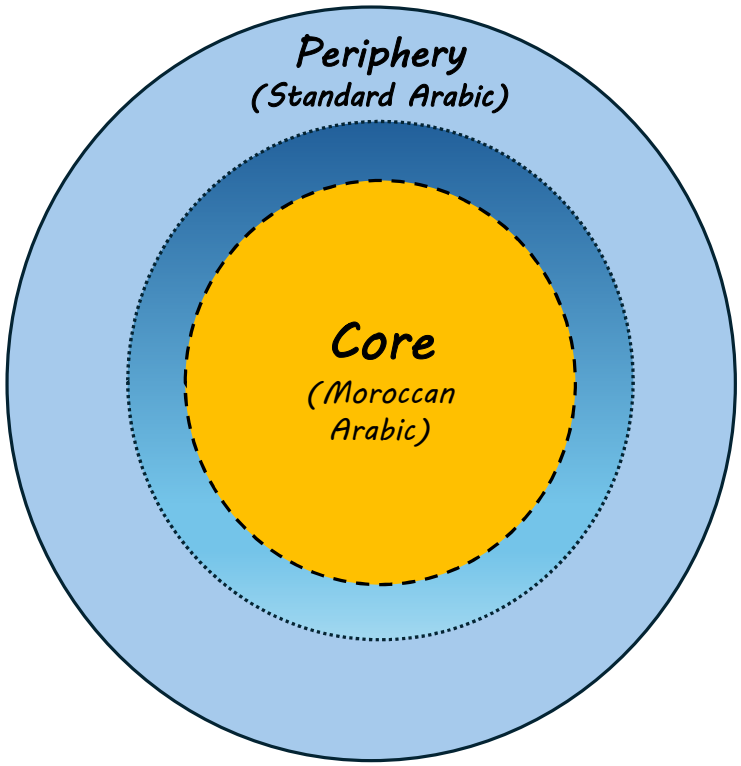
l-ž assimilation is therefore more phonologically than lexically conditioned.

- We argue that neither captures the full picture.
- And propose an alternative model

Our proposal

- We propose a stratified lexicon of MA
- With a **central core** of nativized MA forms
- Surrounded by a (secondary) **periphery** associated with StA.

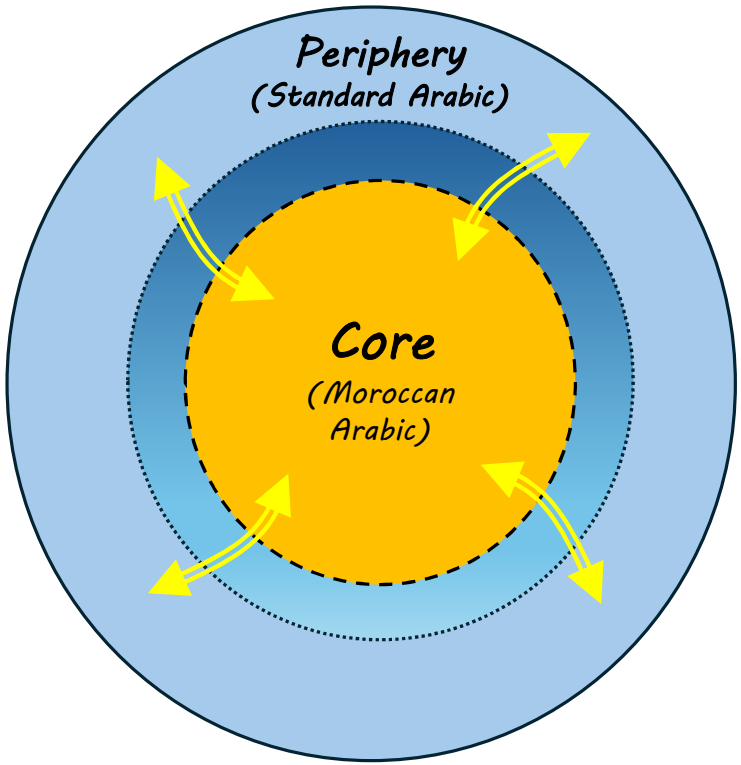
- And a **buffer zone** allowing for **ambiguous** affiliation (dual behavior)



Our proposal

- Crucially, the boundary between the core and periphery is not rigid.

- Rather, t is **permeable**, or porous, allowing movement between periphery and core (dual affiliation)



Under this view

- *l-ž assimilation* is determined by neither phonology nor borrowing status.
- It straightforwardly follows from the architecture of the diglossic mental grammar.
- Specifically,
  - core items assimilate
  - peripheral items resist *l-ž assimilation*
- Interestingly,
  - there is the possibility of items showing ambiguous behavior (i.e., both)

Whence the idea of ‘diglossic grammar’?

ERP Study

- Idrissi et al.'s (2021) ERP study of adjective-noun agreement processing in StA:
  - Full agreement (human nouns)
  - Deflected agreement (non-human nouns)

	Acceptable		**Violation**	
Human	t-tužžaar	l-ʔaθriyyaaʔ	*t-tužžaar	l-θariyy-aat
	masc.pl.	masc.pl.	*masc.pl.	fem.pl.
	'rich merchants'			
Nonhuman	l-ʔusuud	l-žaaʔiʕ-a	*l-ʔusuud	l-žiyyaaʕ
	masc.pl.	fem.sing.	masc.pl.	masc.pl.
	'hungry lions'			

ERP Results

- Human violations (*t-tužžaar l-θariyy-aat)
  - triggered a N400–P600 response
  - indicating they were processed as **genuine grammatical** violations

- Nonhuman violations (of deflected agreement: *l-ʔusuud l-žiyyaaʕ) showed:
  - no significant ERP difference from grammatical sentences
  - suggesting that they were **tolerated** by the parser

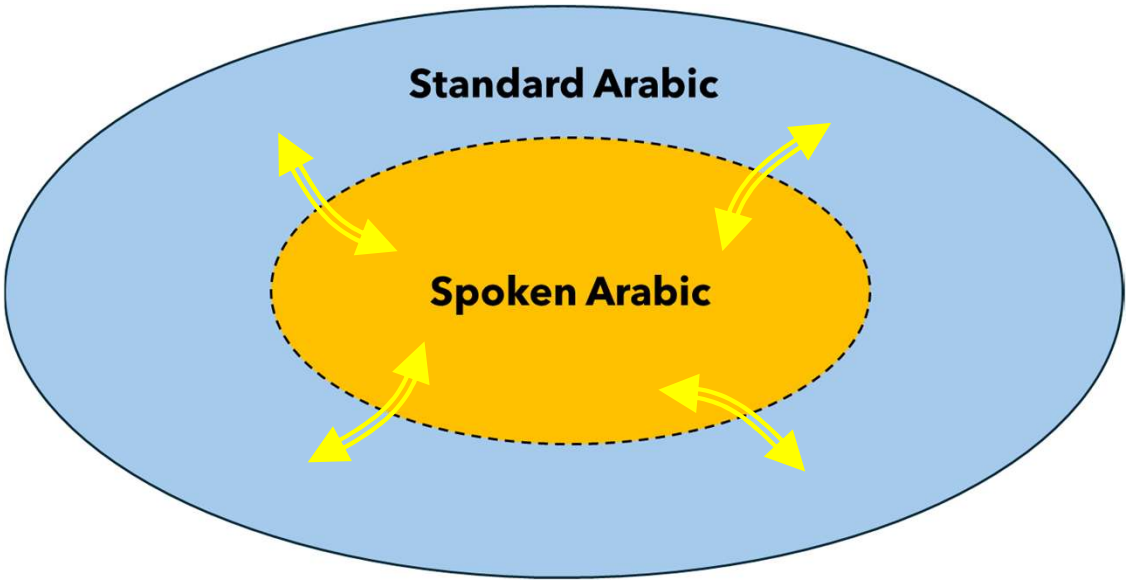
How can these results be explained?

Diglossic grammar

- The answer lies in **diglossia**:
  - StA has strict rules (e.g., deflected agreement is obligatory with nonhuman plurals)
  - Spoken Arabic shows variation (both full and deflected agreement possible)
- Consequently, speakers show **processing flexibility** whereby violations in StA are not necessarily treated as true violations because SpA allows them.

Cognitive consequence

- We argue that Arabic diglossia is characterized by:
- A **hybrid, stratified grammar** where two systems coexist
- The two systems **dynamically interact**



Cognitive consequence

- Because the borders between the two systems are porous, we predict:
  - **Flexible parsing** system, where the parser accesses "loose rules"
  - **Gradient acceptability**: some structures are neither fully grammatical nor fully ungrammatical
- Interactive processing allows **simultaneous** access to both systems

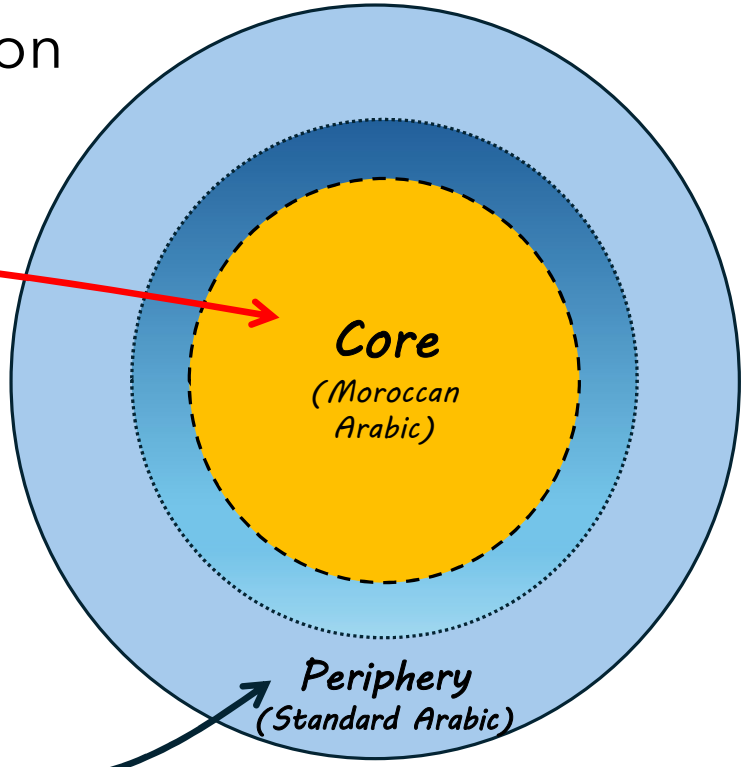
Beyond sociolinguistics

- Diglossia in Arabic is therefore not only a sociolinguistic situation.
- It is cognitively real and determines the architecture of the mental grammar.
- This core/periphery + permeability model inspired our analysis of the MA assimilation puzzle.

The analysis

- Stems in the core trigger l-ž assimilation
- Stems in the periphery do not.

Moroccan Arabic			
/l-žməl/	[žžməl]	*[lžməl]	‘camel’
/l-žərba/	[žžərba]	*[lžərba]	‘scabies’
/l-žanuub/	*[žžanub]	[lžanub]	‘south’
/l-žism/	*[žžism]	[lžism]	‘body’

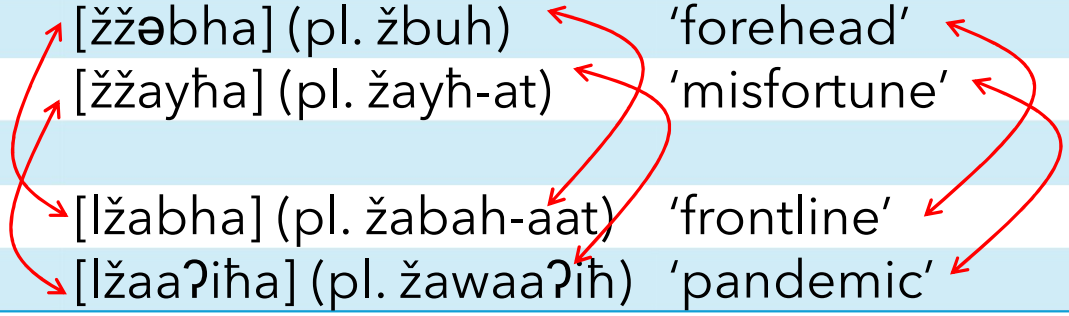


This captures the basic data.

Some independence evidence

- **Near cognates** = **same item but different affiliation**
- Item affiliation coincides with:
  - different assimilation behavior (žž vs. lž)
  - different morphology (broken vs. sound plural)
  - as well as **semantic** differences

Core items (MA)		
/l-žəbha/	[žžəbha] (pl. žbuh)	‘forehead’
/l-žayħa/	[žžayħa] (pl. žayħ-at)	‘misfortune’
Periphery items (StA)		
/l-žabha/	[lžabha] (pl. žabah-aat)	‘frontline’
/l-žaaʔiħa/	[lžaaʔiħa] (pl. žawaaʔiħ)	‘pandemic’



Phonemic inventory

- MA core vocabulary lacks the glottal stop:
  - StA /muʔmin/ ‘believer’
  - MA /mumən/ ‘believer’

Syntax

- Differences extend to syntactic properties.
- Periphery (StA) **participles** avoid the MA circumfixal negative marker *ma–š*
- And select the particle *maši*
- Example: *maši ħaaʔir* ‘not confused’

  **ma-ħaʔir-š*
- Unlike core that takes both:

  ma-xayəf-š ‘not scared’

  maši xayəf

Borrowings

- Notably, MA borrowings tend to get admitted directly into the core:
  - French *jaquette* > /l-žakita/ = [žžakita] ‘jacket’
  - French *journal* > /l-žurnal/ > *žžurnal* ‘newspaper’

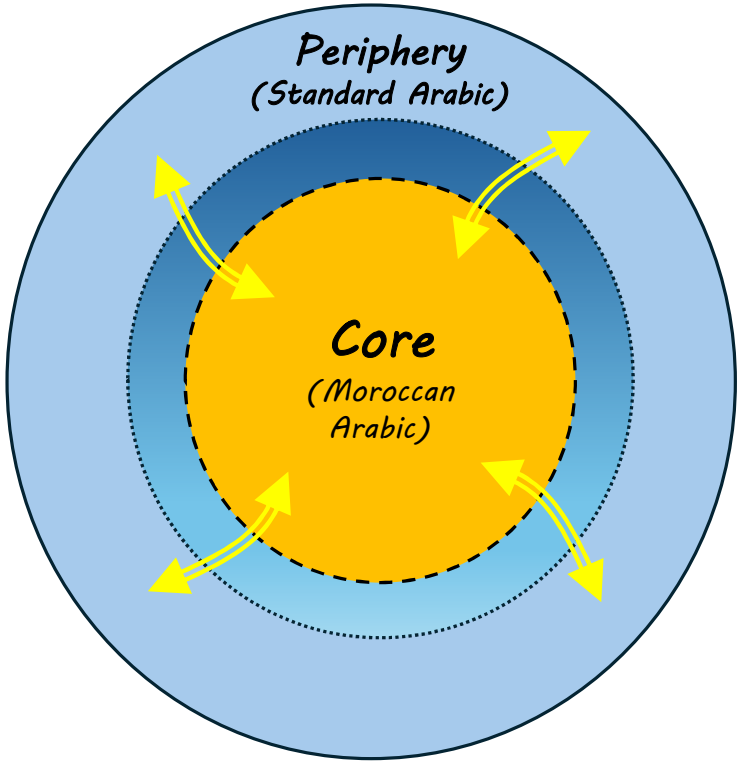
- Prediction: StA borrowings (should) adhere to periphery phonology (need data!)

Ambiguous behavior

- The **appeal** of the proposed model!
- The model predicts dual/ambiguous behavior whereby items can shift between core and periphery depending on extragrammatical factors such as:
  - speaker literacy level
  - register or context
  - language attitudes
  - age
  - gender
  - or subdialect

Ambiguous behavior

- Examples:
  - Arabic /l-žil/ ‘generation’ can be *lžil* or *žžil*
  - French /l-žinirik/ ‘credits’ can be *lžinirik* or *žžinirik*
- This reflects:
  - the interaction between the two systems
  - the porous boundaries
  - and the reality of a buffer zone



Conclusions

- MA l-assimilation data can be accounted for within a model of diglossic grammar where the core and periphery form separate but converging and interacting strata.
- The model explains **l-ž assimilation variability**
- Extends beyond phonology to capture patterns in ***morphology***, ***semantics,*** and **syntax.**
- Links to previous ERP agreement processing data.
- It explains both categorical and gradient patterns.
- Capture ambiguous behavior cases.
- Offers ground for future experimentation!

Thank you.

References

- Becker, M., & Gouskova, M. (2016). Source-oriented generalizations as grammar inference in Russian vowel deletion. *Linguistic Inquiry*, *47*(3), 391–425.
- Freeman, A. (2016). A historical and dialectological perspective: Arabic ǰ and the class of Sun Letters. In S. Davis & U. Soltan (Eds.), *Perspectives on Arabic Linguistics XXVII: Papers from the Annual Symposium on Arabic Linguistics, Bloomington, Indiana, 2013* (pp. 171–185). John Benjamins Publishing Company. https://doi.org/10.1075/sal.3.08fre
- Harrell, R. S. (1962). *A Short Reference Grammar of Moroccan Arabic*. Georgetown University Press (Foreign Service Institute, Arabic Series).
- Heath, J. (1987). *Ablaut and ambiguity: Phonology of a Moroccan Arabic dialect*. State University of New York.
- Hsu, B., & Jesney, K. (2018). Weighted scalar constraints capture the typology of loanword adaptation. In *Proceedings of the Annual Meetings on Phonology 2017: Vol. 5(0)* (pp. 1–13).
- Idrissi, A., Mustafawi, E., Khwaileh, T., & Muralikrishnan, R. (2021). A neurophysiological study of noun-adjective agreement in Arabic: The impact of animacy and diglossia on the dynamics of language processing. *Journal of Neurolinguistics*, *58*, 100964. https://doi.org/10.1016/j.jneuroling.2020.100964
- Itô, J., & Mester, A. (1995). Japanese Phonology. In J. Goldsmith (Ed.), *The Handbook of Phonological Theory* (pp. 817–838). Blackwell.
- Itô, J., & Mester, A. (1999). The structure of the phonological lexicon. In N. Tsujimura (Ed.), *Handbook of Japanese Linguistics* (pp. 62–100). Blackwell.
- Jurgec, P. (2010). Disjunctive lexical stratification. *Linguistic Inquiry*, *41*(1), 149–161.
- Nirheche, A. (2025). Variable and Exceptional Assimilation of the Definite Article [l-] in Moroccan Arabic. In *Proceedings of the Annual Meetings on Phonology: Vol. 1(1)*. https://doi.org/10.7275/amphonology.3041
- Smith, J. L. (2018). Stratified faithfulness in Harmonic Grammar and emergent core–periphery structure. In R. Bennett, A. Angeles, A. Brasoveanu, D. Buckley, N. Kalivoda, S. Kawahara, G. McGuire, & J. Padgett (Eds.), *Hana-bana: A Festschrift for Junko Itô and Armin Mester* (pp. 13–22). University of California.