Contrast dispersion and the positional typology of geminates

Olga Dmitrieva

1 Introduction

(1) Typology: geminate consonants are not equally distributed over various segmental contexts and prosodic positions:

- Typically intervocalic (Kraehenmann, 2001; Thurgood, 1993)
- Avoided at the word-edges (although word-initial to a lesser degree) (Muller, 2001)
- Avoided in adjacency to other consonants (Thurgood, 1993)

(2) These tendencies manifest themselves in

- Morphophonemic alternations
- Dialectal variation
- Historical sound change
- Patterns in segment inventories

(3) **Morphophonemic alternations**

- Degemination in consonant-adjacent context
  Syrian Arabic: [waʔʔef] 'stop' (masc.) - [waʔfi] 'stop' (fem.) (Cowell, 1964)
  Hungarian: [hɔll] - [hɔlvɔ] 'hearing' (Pycha, 2010)
- Degemination in word-final position
  Iraqi Arabic: [maħallaat] 'places' - [maħal] 'place' (Erwin, 1963)
  Maltese: [fomm-i] 'my mouth' - [fom] 'mouth' (Borg, 1997)

(4) **Variation**

- Frequent degemination in consonant-adjacent position
  Russian: [rʌssol] 'brine' - [rʌs(s)kaz] 'story'
- and word-final position
  Russian: [stressa] 'stress' (Gen.) - [stres] 'stress' (Nom.)

(5) **Sound change**

- Word-final degemination
  Ǧubb'adīn dialect of Neo-Arameic: historical neutralization of word-final geminates (Jastrow, 1997)

(6) **Inventories**

- Many languages impose a categorical prohibition against geminate-singleton contrast in certain environments:
  Bengali: only intervocalic geminates
  Cypriot Greek: only intervocalic and word-initial geminates

(7) A near-universal implicational hierarchy:

$$\boxed{\textit{Intervocalic} > \textit{word-initial} > \textit{preconsonantal} > \textit{word-final}}$$

where contrast in a given position entails contrast in positions to the left

(8) Explanation?

- Contextual restrictions are often connected to the relative availability of the phonetic cues to contrast
- Geminate-singleton contrast relies on the durational cues: how much longer is the long?
- Pajak (2009): geminate-singleton ratio in Moroccan Arabic
  $\textit{Initial}+C < \textit{medial}+C < \textit{initial}+V < \textit{medial}+V$
- Ridouane (2007): geminate-singleton ratio in Tashelhit Berber
  $\textit{Intervocalic} < \textit{Initial} < \textit{Final}$

2 Experiment

(9) Beyond the ratio:
Is there still a context-dependent difference in the perceptibility of the contrast?

Experiment: Perception of consonant length by speakers of Russian, American-English, and Italian.
Investigating the effect of

- *Segmental environment*: consonant-adjacent vs. intervocalic (*isek* - *islek*)
- *Word-position*: word-initial vs. word-final (*pos avap* - *po savap*)

Evaluation of the contrast perceptibility: β-coefficients of the logistic function fitted to the individual identification curves.

(10) **Results**
Contrast between short and long consonants is perceived in a more *categorical* fashion in the intervocalic and word-initial than in the preconsonantal and word-final context.

- Steeper identification curves
- A better defined perceptual boundary between two categories

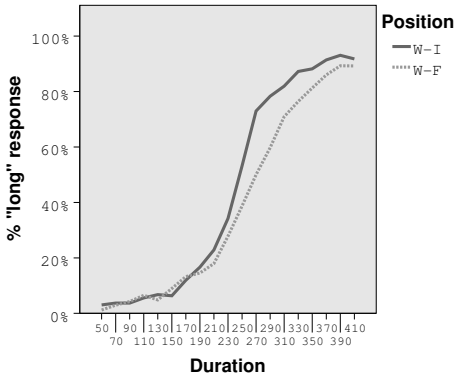

Figure 1: Perception of length in word-initial and word-final consonants.

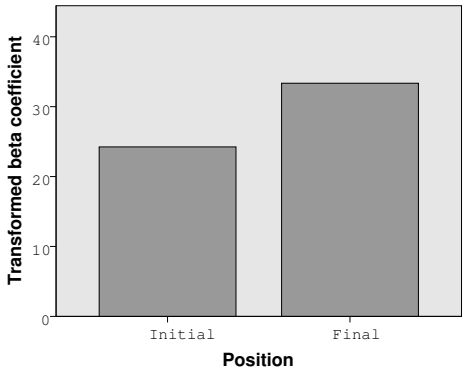

Figure 2: Transformed β-coefficients for word-initial and word-final consonants.

(11) The order of contrast distinctiveness:

Intervocalic > *word-initial* > *preconsonantal* > *word-final*

Context	b coefficient
Intervocalic (V_V)	22.53
Word-initial (WI)	24.24
Preconsonantal (_C)	28.23
Word-final (WF)	33.35

Table 1: Transformed β-coefficients

3 Modeling

(12) **OT Model**

- How do the asymmetries in typology arise from the contextually-driven differences in perceptibility of the contrast?
- Building on The Dispersion Theory of Contrast (Flemming, 1995, 2004)
  - Maximize perceptual contrast distinctiveness
  - Maximize number of contrasts
  - Minimize articulatory effort

(13) Perceptually-motivated constraints on distinctiveness of the durational contrast in consonants:

- MinDist$_n$ = Maintain contrast distinctiveness at the degree n.
  - MinDist$_{V_V}$ = Maintain contrast distinctiveness at the intervocalic level.

- MinDist$_{WI}$ = Maintain contrast distinctiveness at the word-initial level.
- MinDist$_{_C}$ = Maintain contrast distinctiveness at the preconsonantal level.

(14) Contrast maximizing and effort minimizing constraints:

- MaxContrast = Maintain a contrast along a given phonological dimension
  - MaxContrast$_L$ = Maintain a length contrast.
- LAZY$_n$ = Do not expend effort of the degree n.
  - LAZY$_L$ = Do not lengthen consonants.

(15)

	$MD_{\_C}$	MD_{WI}	MD_{V_V}	$MaxC$	$Lazy$
☞1a. ata-atta					*
1b. ata				*!	
1c. atta				*!	*
2a. ta-tta			*!		*
☞2b. ta				*	
2c. tta				*	*!
3a. atCa-attCa		*!	*		*
☞3b. atCa				*	
3c. attCa				*	*!
4a. at-att	*!	*	*		*
☞4b. at				*	
4c. att				*	*!

(16) The factorial typology of this model contains five language types:

Output 1	Output 2	Output 3	Output 4	Output 5
ata	ata-atta	ata-atta	ata-atta	ata-atta
ta	ta	ta-tta	ta-tta	ta-tta
atCa	atCa	atCa	atCa-attCa	atCa-attCa
at	at	at	at	at-att

Table 2: The factorial typology.

(17) Implicational relationships within the factorial typology:

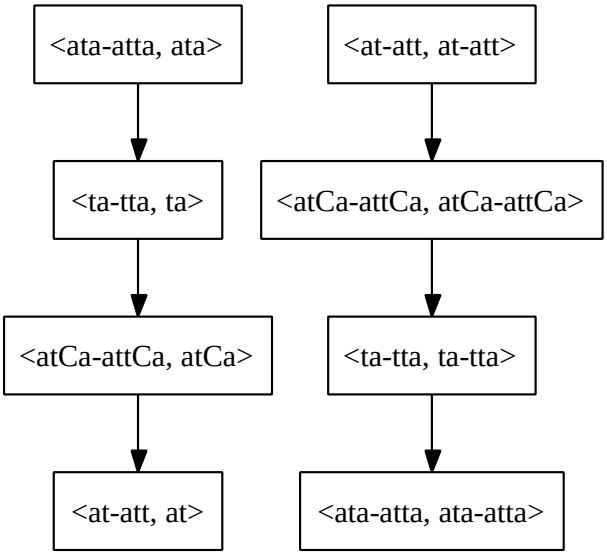

Figure 3: Implicational universals in the factorial typology

(18)

	Grammar	Language	Observed frequency
Type 1	No geminates	English Spanish Mandarin	95%
Type 2	Intervocalic geminates	Amharic Bengali Madurese Ge'ez Maranguku Oromo Somali Yagua	2%
Type 3	Intervocalic and W-I geminates	Cypriot Greek Iraqi Arabic Syrian Arabic Selayarese Finnish	1.25%
Type 4	Intervocalic, W-I, and C-adjacent geminates	Syrian Arabic Maltese	0.8%
Type 5	Intervocalic, W-I, C-adjacent and W-F geminates	Moroccan Arabic Berber	0.8%

Table 3: Five language types predicted by the model.

(19) Frequency distribution:
No geminates >> V_V > V_V+WI > no WF, all

(20) Multiple Grammars + ranking volume approach to modeling frequency:

	Grammar	Ranking volume
1	No geminates	60
2	Intervocalic geminates	20
3	Intervocalic and W-I geminates	10
4	Intervocalic, W-I, and C-adjacent geminates	6
5	Intervocalic, W-I, C-adjacent and W-F geminates	24

Table 4: Number of total orders (ranking volume) that derive each output pattern.

- Problem: Language type 5 is predicted to be rather common.

(21) The Initiality exception: languages like Pattani Malay allow word-initial but not intervocalic geminates.

- Contrast preservation in the psycholinguistically salient position:
- MaxContrast_{WI} = Maintain contrast word-initially.

(22)

	$MaxContrast_{WI}$	$Lazy$
1a. ata-atta		*!
☞1b. ata		
1c. atta		*!
☞2a. ta-tta		*
2b. ta	*!	
2c. tta	*!	*
3a. atCa-attCa		*!
☞3b. atCa		
3c. attCa		*!
4a. at-att		*!
☞4b. at		
4c. att		*!

(23)
- An additional language type is generated
- But the implicational relationship between word-initial and intervocalic contrasts is destroyed.

Output 1	Output 2	Output 3	Output 4	Output 5	Output 6
ata	ata-atta	ata-atta	ata-atta	ata-atta	ata
ta	ta	ta-tta	ta-tta	ta-tta	ta-tta
atCa	atCa	atCa	atCa-attCa	atCa-attCa	atCa
at	at	at	at	at-att	ta

Table 5: The factorial typology with the Initiality constraint

(24) The morphological factor:
- True and concatenated geminates are phonetically equivalent
- The same contrast-distinctiveness constraints are expected to apply to both.
- Additional evidence that concatenated geminates are realized even if true geminates are not.
  - *Hungarian*: true _C → degemination, concatenated can be preserved
  - *Russian*: concatenated geminates are less susceptible to degemination.
- MaxContrast$_M$ = Maintain a contrast between morphemes

(25) *Russian*: *po-dat^j* - *pod-dat^j*
- In case of degemination, the contrast between the morphemes is neutralized:
- **po**-*dat^j* - **po**-*dat^j*

(26) A model with this additional constraint generates a factorial typology of 19 languages.
- The same implicational relationship hold for true and concatenated geminates: *Intervocalic* > *word-initial* > *preconsonantal* > *word-final*
- Additional implication:

$$\boxed{\textit{concatenated} > \textit{true}}$$

contrastive consonant length morpheme-internally implies preservation of length in the concatenated geminates in the same position.

(27) Implicational relationships in the model are shown in Figure 4.

(28) The Enhancement alternative: for a complete treatment of the contrast typology as function of the contextually-driven differences in perceptibility we need to consider *contrast enhancement* as an alternative to neutralization in perceptually disadvantaged positions.
- *Maltese*: vowel epenthesis before geminate-initial words

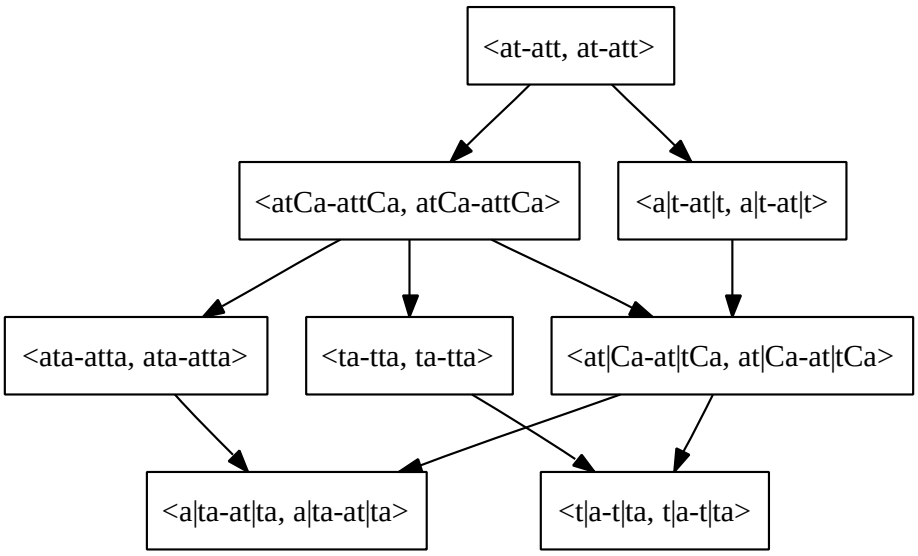

Figure 4: Implicational universals of the model with morphological constraint.

- *Pattani Malay*: initial syllables with geminated onsets - higher amplitude and fundamental frequency than initial syllables with singleton onsets
- *Cypriot Greek*: both word-initial and intervocalic geminates are aspirated.

(29) Although difficult to verify empirically, it is reasonable to expect enhancement to apply more readily to contrasts with jeopardized perceptibility.
It is also likely that enhancement is produced at the expense of additional articulatory effort.

- Taking aspiration as an example of enhancement:
- LAZY$_A$ = "do not aspirate consonants"

(30)

	$MD_{\_C}$	MD_{WI}	MD_{V_V}	$MaxC$	$MaxC_{WI}$	$Lazy_L$	$Lazy_A$
☞1a. ata-atta						*	
1b. ata-att^{h}a						*	*!
1c. ata				*!			
1d. atta				*!		*	
2a. ta-tta			*!			*	
☞2b. ta-tt^{h}a						*	*
2c. ta				*!	*		
2d. tta				*!	*	*	
3a. atCa-attCa		*!	*			*	
3b. atCa-att^{h}Ca			*!			*	*
☞3c. atCa				*			
3d. attCa				*		*!	
4a. at-att	*!	*	*			*	
4b. at-att^h		*!	*			*	*
☞4c. at				*			
4d. att				*		*!	

(31) The factorial typology of the enhancement model includes 17 languages shown in Table 6.

- Outputs 1-6 the basic patterns we saw above
- Outputs 7-13 demonstrate follow implicational hierarchy:

$$\boxed{\textit{word-final}^h > \textit{preconsonantal}^h > \textit{word-initial}^h}$$

where contrast enhancement in a given position entails contrast enhancement in positions to the left

- Outputs 14-17 are abnormal in that they do not follow this implicational pattern
- Although 15 and 16 can be expected under the assumption that enhancement of the contrasts is desirable in psycholinguistically salient positions.

(32) This model does not generate Cypriot Greek, where contrast is present intervocalically and word-initially *and enhanced in both contexts.*

- Such patterns may arise from constraint on phonetic uniformity of the phonological contrast across contextual environments:
- Unify$_P$ = Contextual paradigms where the same phonological contrast varies in its phonetic realization are prohibited.

Output 1	Output 2	Output 3	Output 4	Output 5	Output 6
ata	ata-atta	ata-atta	ata-atta	ata-atta	ata
ta	ta	ta-tta	ta-tta	ta-tta	ta-tta
atCa	atCa	atCa	atCa-attCa	atCa-attCa	atCa
at	at	at	at	at-att	ta

Output 7	Output 8	Output 9	Output 10	Output 11	Output 12	Output 13
ata-atta	ata-atta	ata-atta	ata-atta	ata	ata-atta	ata-atta
ta-tta	ta-tta	ta-tt^{h}a	ta-tt^{h}a	ta-tt^{h}a	ta-tta	ta-tt^{h}a
atCa-attCa	atCa-att^{h}Ca	atCa-att^{h}Ca	atCa	atCa	atCa-att^{h}Ca	atCa-att^{h}Ca
at-att^h	at-att^h	at-att^h	at	at	at	at

Output 14	Output 15	Output 16	Output 17
ata-atta	ata-atta	ata-atta	ata-atta
ta-tta	ta-tt^{h}a	ta-tt^{h}a	ta-tt^{h}a
atCa-at^{h}Ca	atCa-attCa	atCa-attCa	atCa-att^{h}Ca
at-att	at-att	at-att^h	at-att

Table 6: The factorial typology of the enhancement model

(33) Addition of the uniformity constraints allows the model to generate 3 additional language types, where the enhancement spreads throughout the contextual paradigm of the phonological contrast.

Output 18	Output 19	Output 20
ata-att^{h}a	ata-att^{h}a	ata-att^{h}a
ta-tt^{h}a	ta-tt^{h}a	ta-tt^{h}a
atCa-at^{h}Ca	atCa-att^{h}Ca	atCa
at-att^h	at	at

Table 7: The factorial typology with uniformity constraint

(34) **Conclusions**

- Perceptually-based constraints on contrast distinctiveness → implicational hierarchy → most of the languages with geminate consonants.
- Aa unified account of the distributional typology of geminate consonants in diverse languages.
- Initially constraint → languages with initial geminate-singleton contrast only.
- Initially constraint → bad quantitative prediction: frequency of word-initial geminates = frequency intervocalic geminates.
- Morphological factor → concatenated geminates are more common than lexical ones.

- Typology of lexical and concatenated geminates: e.g. no language with lexical geminates but not concatenated geminates in a particular context.
- Enhancement alternative → typology of durational contrast enhancement (yet to be verified empirically): e.g. languages with enhanced intervocalic contrast must enhance all others.

4 *

References

Borg, A. (1997). Maltese phonology. In Kaye, A. S., editor, *Phonologies of Asia and Africa*, pages 245–285. Eisenbrauns, Winona Lake, Indiana.

Cowell, M. (1964). *A reference grammar of Syrian Arabic. Based on the dialect of Damascus.* Georgetown University Press, Washington, D.C.

Erwin, W. M. (1963). *A short reference grammar of Iraqi Arabic.* Georgetown University Press, Washington, D.C.

Flemming, E. (1995). *Auditory representations in phonology.* PhD thesis, UCLA.

Flemming, E. (2004). Contrast and perceptual distinctiveness. In Hayes, B. Kirchner, R. and Steriade, D., editors, *Phonetically Based Phonology*, pages 232–276. Cambridge Univesity Press, Cambridge.

Jastrow, O. (1997). The Neo-Aramaic languages. In Hetzron, R., editor, *The Semitic languages*, pages 312–377. Routledge, New York.

Kraehenmann, A. (2001). Swiss German stops: Geminates all over the word. *Phonology*, 18:109–145.

Muller, J. (2001). *The Phonology and Phonetics of Word-Initial Geminates.* PhD thesis, Ohio State University.

Pajak, B. (2009). Context-dependent perception of geminates. Poster presented at the 83rd Annual Meeting of the Linguistic Society of America, San Francisco.

Pycha, A. (2010). A test case for he phonetics-phonology interface: gemination restrictions in Hungarian. *Phonology*, 27:119–152.

Ridouane, R. (2007). Gemination in Tashlhiyt Berber: an acoustic and articulatory study. *Journal of the International Phonetic Association*, 37:119–142.

Thurgood, G. (1993). Geminates: a cross-linguistic examination. In Nevis, J. A., M. G. and Thurgood, G., editors, *Papers in honor of Frederick H. Brengelman on the occasion of the twenty-fifth anniversary of the Department of Linguistics, (CSU) Fresno*, pages 129–139. Fresno, CA.