

A segment-specific metric for quantifying participation in harmony



Avery Ozburn
Department of Linguistics, University of British Columbia

Gradient degrees of participation/neutrality in harmony are quantified in three languages

Summary

New measure for segment-specific harmony participation
Possible correlations with categorical neutrality patterns

Key questions

How can we measure a segment's participation in harmony?
Do gradient patterns of harmony participation mirror cross-linguistic trends for categorical participation?

Introduction

Typological patterns and variability in participation
Might expect to see patterns gradiently
No systematic way of calculating the degree to which a segment participates in a harmony system

Relative risk (RR)

Common statistical measure
Relative probability of an event in one condition versus another
Here, interpreted as:
Event = specific vowel
Condition = harmonic vs. non-harmonic contexts
Provides frequency-correct measure of probability of vowel in harmony contexts
i.e. how harmonic the vowel is
Previously used in linguistics for speech errors (Tupper and Alderete in prep)
RR for vowel V from category C1 can be calculated as in (1):

$$(1) RR(V) = \frac{(\# V \text{ in context } C1) \times (\# \text{ anything in context } C2)}{(\# V \text{ in context } C2) \times (\# \text{ anything in context } C1)}$$

Two RR(V)s for each V – V1 position (preceding) and V2 position (following)

Motivations

Recent trend of investigating gradient co-occurrence patterns in harmony systems (e.g. Pierrehumbert 1993, Frisch et al. 2004),
Segment-specific participation is complex, could benefit from treatment as gradient
Certain vowels are typologically favoured as triggers and targets
Non-participation seems to be correlated with specific vowel qualities (e.g. Mayak)
Might expect these patterns on a gradient level
Variation and idiosyncratic behaviour in harmony
Neutrality and participation are not clear-cut categories (e.g. Hungarian)
Categorical descriptions and existing measures (O/E, harmony index, % harmonic) cannot capture potentially interesting segment-specific gradient trends

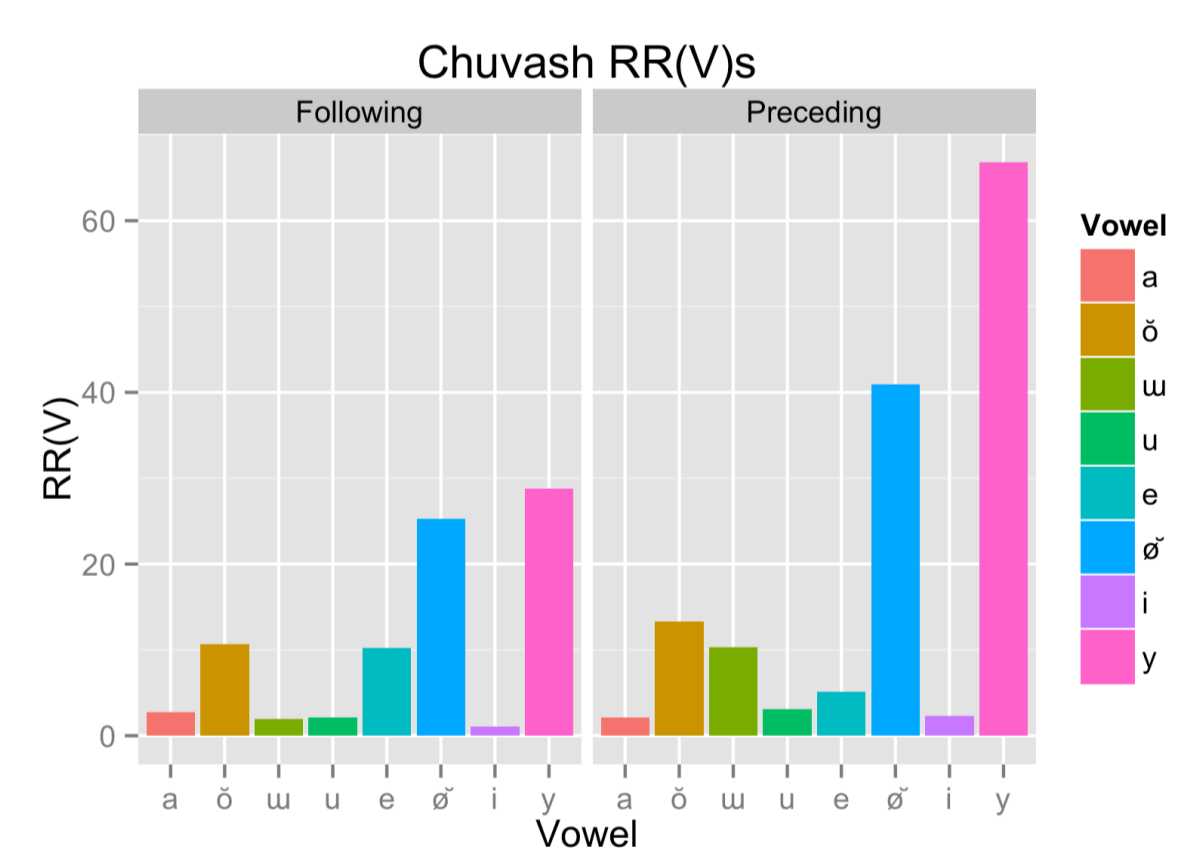
Corpora background

Front/back harmony in 3 corpora
Chuvash, Tatar, and Mari
Luutonen et al. 2007; Luutonen et al. 2016
Dictionary lists
Include derived forms
Broad harmonic behaviour through lexicon
Counts of pairs of syllable-adjacent vowels
Using Phonological CorpusTools
O/E values for all vowel pairs
RR(V) for each vowel

Chuvash

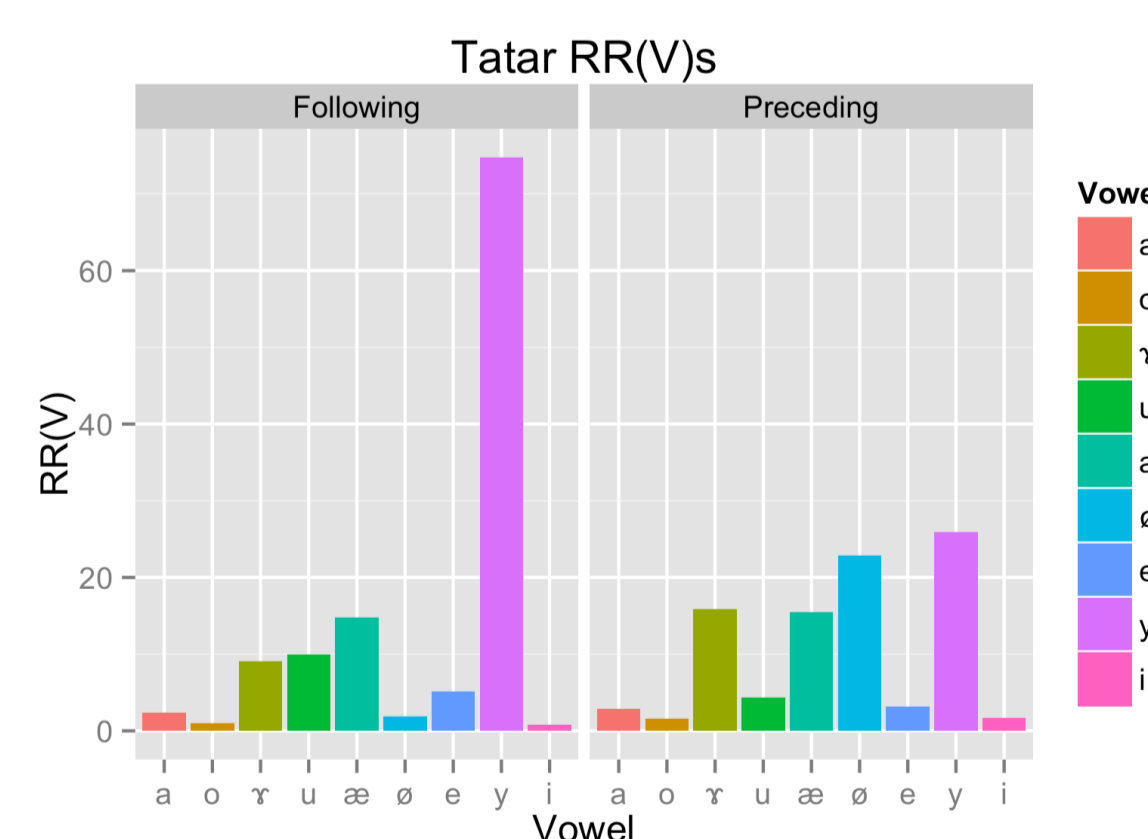
	a	ö	u	u	e	ö	i	y
a	1.30	1.14	1.36	1.54	0.32	0.08	1.49	0.06
ö	1.41	2.03	0.94	0.84	0.02	0.11	0.27	0.09
u	1.26	2.22	0.34	0.89	0.09	0.02	0.35	0.10
u	1.23	1.62	1.56	1.20	0.31	0.14	0.92	0.15
e	0.48	0.15	0.52	0.61	2.26	1.64	1.55	3.98
ö	0.04	0.04	0.62	0.26	2.70	4.43	0.24	3.18
i	0.88	0.24	0.86	0.91	1.79	1.49	1.21	0.45
y	0.03	0.02	0.00	0.17	3.03	4.24	0.15	1.95

Results



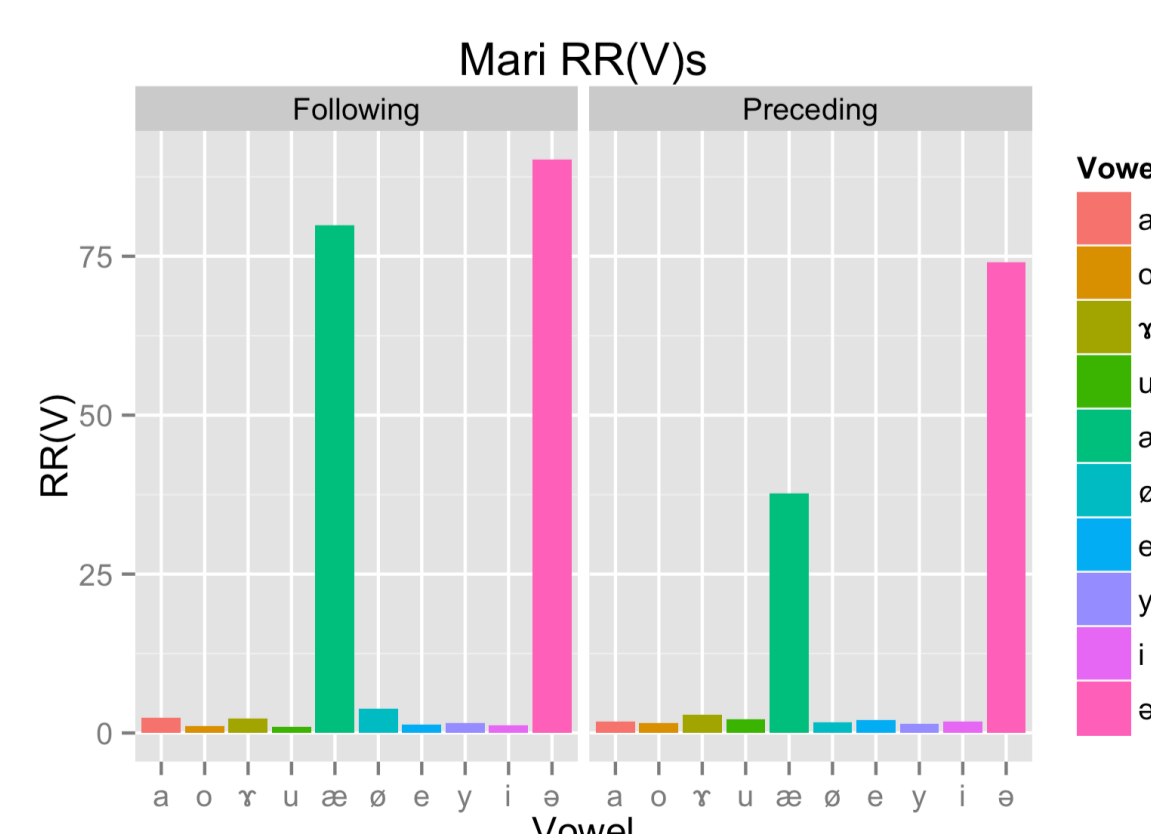
Tatar

	a	o	ɣ	u	æ	ø	e	y	i
a	1.24	0.90	1.48	2.01	0.21	0.49	0.31	0.03	1.29
o	1.32	2.60	0.94	0.42	0.04	0.46	0.82	0.03	1.80
ɣ	1.27	0.18	2.50	2.27	0.09	1.22	0.14	0.04	0.08
y	1.94	0.44	1.76	0.61	0.08	1.23	0.32	0.03	0.84
æ	0.23	0.03	0.08	0.07	2.68	0.86	1.80	4.75	0.63
ø	0.17	0.08	0.03	0.04	3.40	3.98	2.97	0.71	0.30
e	0.61	1.34	0.27	0.21	1.52	2.17	1.88	2.12	1.09
y	0.12	0.08	0.06	0.06	4.23	1.48	2.49	0.59	0.31
i	1.04	2.05	0.24	0.27	1.50	0.39	1.43	0.57	1.23



Mari

	a	o	u	ɣ	æ	ø	ɣ	e	i	ə
a	1.32	0.71	1.17	1.12	0.06	0.16	0.93	0.85	1.26	0.05
o	0.96	2.05	0.96	1.09	0.04	0.20	0.64	1.05	1.43	0.02
u	1.01	1.50	1.36	1.38	0.04	0.37	0.77	0.84	0.99	0.04
ɣ	1.41	0.60	0.63	1.37	0.01	1.20	0.88	0.92	0.26	0.02
æ	0.06	0.10	0.24	0.08	8.46	0.07	0.84	0.26	0.27	6.14
ø	0.72	0.27	0.60	1.16	1.64	13.90	1.72	1.01	0.25	1.81
y	0.68	0.37	0.66	1.44	1.15	9.32	4.37	1.11	0.13	1.52
e	0.71	1.09	1.19	0.62	1.49	0.39	1.34	1.60	1.78	1.29
i	0.67	1.84	1.93	0.46	1.55	0.44	0.58	1.40	1.69	1.22
ə	0.02	0.11	0.18	0.02	6.50	0.43	0.60	0.60	0.12	8.63



Discussion

Correspondence between RR(V) and informal impression of how much V participates in harmony, based on O/E
Chuvash [y]: O/E 0 to 0.17; RR(y) is very high (66.85)
Chuvash [i]: O/E 0.24 to 0.91; RR(i) is low (2.30)
Chuvash [u]: O/E 0.02 to 0.35; RR(u)=10.31
In all three languages, RR(i) is very low
[i] has a low degree of participation in harmony
[i] is a very common transparent vowel
Gradient non-participation here
Cardinal vowels tend to have lower RR(V) – i.e. less harmonic
Non-harmonic Russian loans?
Influence of frequency in morphology?

Future directions

More corpora!
Other languages with front/back harmony
Particularly with categorical neutrality
Other types of harmony
Non-harmony languages

Acknowledgements

Thank you very much to Paul Tupper for suggesting possible metrics to use for this question, and to Gunnar Hansson and Doug Pulleyblank for feedback.

References

- Agresti, Alan. 2013. Categorical data analysis. John Wiley & Sons.
Andersen, Torben. 1999. Vowel harmony and vowel alternation in Mayak (Western Nilotic). *Studies in African Linguistics*, 28(1), 1-30.
Frisch, Stefan A., Janet B. Pierrehumbert, & Michael B. Broe. 2004. Similarity avoidance and the OCP. *Natural Language and Linguistic Theory* 22: 179-228.
Hall, K. C., Allen, B., Fry, M., Mackie, S., and McAuliffe M. 2015. Phonological CorpusTools, Version 1.1.0. [Computer program]. Available from <https://sourceforge.net/projects/phonologicalcorpusools/>.
Kaun, Abigail R. 1995. The typology of rounding harmony: an optimality theoretic approach. Los Angeles, CA: UCLA dissertation.
Kiparsky, Paul & Karl Pajusalu. 2003. Toward a typology of disharmony. *The Linguistic Review*, 20, 217-241. DOI: 0167-6318/03/020-0217
Luutonen, Jorma, et al. (ed.). 2007. *Electronic Word Lists: Mari, Mordvin and Udmurt*. With SFou WordListTool 1.3. Lexica Societatis Fenno-Ugricae XXXI:1. ISBN 978-952-5150-98-8. Helsinki.
Luutonen, Jorma, et al. (ed.). 2016. *Electronic Word Lists: Komi, Chuvash and Tatar*. With SFou WordListTool 1.4. Lexica Societatis Fenno-Ugricae XXXI:2. ISBN 978-952-5667-79-0. Helsinki.
Pierrehumbert, Janet B. 1993. Dissimilarity in the Arabic verbal roots. *Proceedings of NELS 23*: 367-381.
Siptár, Péter, and Miklós Törkenczy. 2000. *The phonology of Hungarian*. Oxford: Oxford University Press.
Tupper, Paul, and John Alderete. In prep. Risk ratios for linguists.