

Phrasal homophony: an interaction between syntax and phonology

Erin Zaroukian
zaroukian@cogsci.jhu.edu
GLS 2010

1 Overview

- Data - phonological repetition avoidance, but not as you've seen it before
 - certain homophonous phrases cannot appear adjacently in Russian
- OT analysis
 - This strange data can be given a fairly standard and unifying account
- Discussion
 - Is this the best analysis for this data?

2 Data

- Russian is a multiple *wh*-fronting language, as illustrated in (1), but multiple *wh*-fronting does not occur if it would result in adjacent homophonous *wh*-phrases, as in (2) (Bošković 2002)
 - Bošković chalked it up to PF constraint against adjacent homophonous *wh*-phrases.

- (1) Kto kogo ljubit? (Bošković 2002:354)
who whom loves
'Who loves whom?'
- (2) a. *Čto čto obuslovalo? (Bošković 2002:364)
what what conditioned
b. Čto obuslovalo čto?
what conditioned what
'What conditioned what?'

- It's not just sequences of homophonous *wh*-phrases which are banned¹.
 - In (4), sequences of homophonous nouns appear to be banned.
 - In (3)-(5), *est'*-insertion, i.e. the pronunciation of the typically-null copula, is available as a repair.²
 - In (5), sequences of homophonous noun phrases appear to be banned.

¹I have attempted to make note when relevant, but not all native speakers agree on each data point.

²Some speakers find the sentences to still be ungrammatical with the insertion of *est'* and prefer instead *Vrač - éto vrač* (doctor that-is doctor), etc. The insertion of the emphatic *že* may be another repair strategy.

- (3) a. *Kto kto.
who who
b. Kto est' kto.
who is who
'Who is who?'
- (4) a. *Vrač vrač.
doctor doctor
b. Vrač est' vrač.
doctor is doctor
'The/a doctor is the/a doctor.'
c. On vrač
he doctor
'He is a doctor.'
- (5) a. *Bogatyj vrač bogatyj vrač.
rich doctor rich doctor
b. Bogatyj vrač est' bogatyj vrač.
rich doctor is rich doctor
'The rich doctor is a rich doctor.'

(Partee 1998)

- How far should this constraint go?
 - Not all sequences of homophonous phrases are banned.
 - In (6), homophonous sequences involving adjectives appear to be fine.³
 - In (7), sequences of homophonous verbs appear to be fine.
- (6) a. Bogatyj bogatyj.
rich rich
'The rich person is rich.'
b. Bogatyj bogatyj vrač.
rich rich doctor
'The rich person is a rich doctor.'
c. Vrač bogatyj vrač.
doctor rich doctor
'The doctor is a rich doctor.'
- (7) Emu nraivitsja nraivit'sja.
he like to-like
'He likes to like'

Summary of data

³Some speakers, however, still require an overt copula in (6).

(8)

good	bad	ex.
	<i>wh wh</i>	(2), (3)
	N/DP N/DP	(4)
	DP DP	(5)
A/DP A		(6)
V V		(7)

- This data is unusual
 - Prohibition is not purely phonological (cf. (4) and (6))
 - Prohibition applies to relatively large syntactic units (e.g. (5))

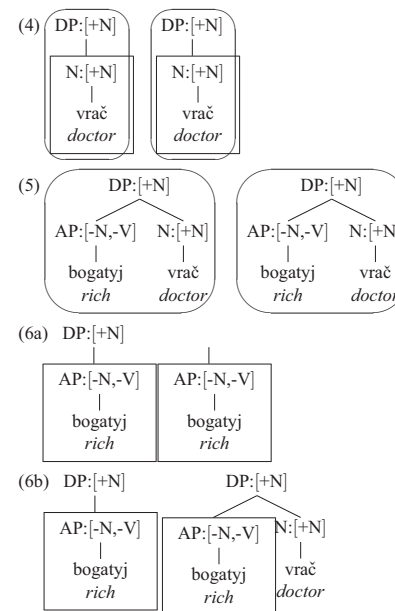
3 An OT analysis

- Obligatory Contour Principle (OCP) - prohibits adjacent identical elements
 - elements \approx phonological, morphemic
 - used e.g. to account for lack of successive occurrences of
 - [+lab] segments in certain languages (e.g. Fukazawa (1999), who uses OCP[lab])
 - homophonous *-s* morphemes in English (e.g. Yip (1998), who uses OCP(s) to explain haplogy in *cats'/*cats's* and *Katzes'/*Katzes's*)
- The data presented resembles OCP, but is different in that it seems to apply at a phrasal (not phonological/morphemic) level, barring homophonous syntactic chunks
- Turns out that this can be given a standard OCP-type analysis
- What differentiates the good from the bad in (8)?

- They are bad only when two nominal elements are involved
- This difference can be captured with category features (Baker 2003)

Noun:	+N	
○ Verb:	+V	
Adjective:	-N,-V	
○ They are bad only when two +N categories are involved		
good	bad	ex.
	<i>wh_[+N] wh_[+N]</i>	(2), (3)
	N _[+N] /DP _[+N] N _[+N] /DP _[+N]	(4)
	DP _[+N] DP _[+N]	(5)
A _[-N,-V] /DP _[+N] A _[-N,-V]		(6)
V _[+V] V _[+V]		(7)

- Data from above, showing adjacent homophonous elements



- Note that in (6a) and (6b) there are no adjacent homophonous [+N] elements, and at the DP:_[+N] level in (6b), the phrases are not homophonous.

- Constraints

- DEP - violated when an element in the output does not have a corresponding element in the input
- OCP(PH) - violated when phrases sharing the same phonological form occur adjacently⁴
- OCP([+N]) - violated when elements sharing the feature [+N] occur adjacently⁵
- (OCP(PH), OCP([+N])) - violated when both OCP(PH) and OCP([+N]) are violated

- The tableau in (9) depicts the contrast in (4), where an overt copula is required.

⁴This constraint is akin to the morphological OCP constraints used in Yip (1998), but it applies to units at a syntactic level.

⁵This constraint is akin to the phonological OCP constraints used widely, but it applies to units at a syntactic level. Also cf. Richards (2006)

- The tableau in (10) depicts the data in (6a), where an overt copula is not required.

(9)		vrač _[+N] vrač _[+N]	(OCP(PH),OCP([+N]))	DEP	OCP(PH)	OCP([+N])
a.		vrač _[+N] vrač _[+N]	*!		*	*
b.		vrač _[+N] est' vrač _[+N]		*		

(10)		bogatyj _[+N] bogatyj _[-N]	(OCP(PH),OCP([+N]))	DEP	OCP(PH)	OCP([+N])
a.		bogatyj _[+N] bogatyj _[-N]			*	
b.		bogatyj _[+N] est' bogatyj _[-N]		*!		

- A violation of OCP([+N]) or OCP(PH) alone is not enough to prompt copula insertion, it is only the conjunction of these constraints, which outranks DEP, that leads to the pronunciation of the copula.
 - It's not all OCP([+N]) - adjacent non-homophonous nouns are common in Russian copular constructions
 - It's not all OCP(PH) - adjacent homophonous items are sometimes allowed, as seen in (6).
- This type of analysis also works for the *wh* data, using *Q (Legendre et al. 1998), which is violated when a *wh*-feature is left unchecked⁶

(11)		čto _[+N] čto _[+N] obuslovalo	(OCP(PH),OCP([+N]))	*Q	OCP(PH)	OCP([+N])
a.		čto _[+N] čto _[+N] obuslovalo	*!		*	*
b.		čto _[+N] obuslovalo čto _[+N]		*		

Summary

- This analysis demonstrates how a phonological/syntactic OCP effect on a phrasal level can be accounted for by defining similarity in terms of features and using familiar OT machinery (cf. Ackema 2001).

4 Discussion

- OT analysis
 - Allows us to account for this data in the same way we can account for similar data
 - Potential problems with an OT analysis
 - Should these types of constraints be used in this way?
 - Constraints like OCP(PH) normally apply to morphological units
 - Constraints like OCP([+N]) normally apply to phonological features

⁶Or, if you believe *wh*-movement in Russian is driven by Focus, this could be reformulated in terms of Focus features.

- Walter (2007) argues that there are three different biases that result in repetition avoidance and that a unified account may not be ideal. The data here would presumably stem from a distinctness condition on linearization (discussed below), and not from physiological or perceptual considerations, which are typically captured by OT.
- Why should we independently believe in OCP([+N])?
 - You do seem to see something like OCP([+N]), e.g. “*This’s cold,*” said *John to Mary*” vs. “*It’s cold,*” told *John Mary* (Richards 2006), though see distinctness condition below
- Should we be mixing syntactic and phonological constraints in the tableaux?
 - It’s not clear (changes the typology), but other people do (e.g. Grimshaw (1997))
- Should category features be used beyond lexical items?
- We haven’t seen anything outrank (OCP(PH),OCP([+N])), so we’re not seeing OT’s strength in explaining that kind of not-across-the-board prohibition (have adjacent homophonous DP rather than violate X)

- Other analyses (see appendices for worked-out examples)

- PF constraint - Bošković
 - PF constraint against consecutive homophonous DPs
 - But this requires access to relevant phonological, syntactic, and linearity information, and it’s not clear that all this information would be simultaneously available at any point in the derivation.
- Distinctness condition - Richards
 - Things that are too similar cannot be linearized (it’s all syntactic!)
 - e.g. In “*It’s cold,*” told *John Mary*, the two DPs (*John* and *Mary*) cannot be linearized; they result in pairs like <DP,DP> (via Kayne (1994)’s LCA) which leads to a crash at PF.
 - It’s not clear how pronunciation of the copula helps, though

5 Summary

- Data showing an OCP-type effect that seems to both syntactic and phonological information
- Unified OT analysis is possible
 - (OCP([+N]),OCP(PH))
- but whether or not a unifying account is desirable is left as an open question.

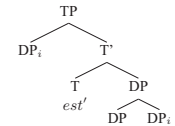
References

- Ackema, P. (2001). Colliding complementizers in Dutch: Another syntactic OCP effect. *Linguistic Inquiry*, 717–727.
- Baker, M. C. (2003). *Lexical categories: verbs, nouns, and adjectives*. Cambridge: Cambridge University Press.
- Bošković, Ž. (2002). On multiple wh-fronting. *Linguistic Inquiry* 33(3), 351–383.
- Franks, S. (1998). Clitics in slavic. Position paper presented at the Comparative Slavic Morphosyntax Workshop.
- Fukazawa, H. (1999). Theoretical implications of OCP effects on features in *Optimality Theory*. Ph. D. thesis, University of Maryland.
- Geist, L. (2008). Predication and equation in copular sentences: Russian vs. English. In I. Comorovski and K. von Heusinger (Eds.), *Existence: Semantics and Syntax*, pp. 79–105. Springer.
- Grimshaw, J. (1997). The best clitic: Constraint interaction in morphosyntax. In L. Haegeman (Ed.), *Elements of Grammar: Handbook of Generative Syntax*, pp. 169–196. Kluwer Academic Publishers.
- Kayne, R. (1994). *The Antisymmetry of Syntax*. Linguistic Inquiry Monograph No. 25. Cambridge: The MIT Press.
- Legendre, G., P. Smolensky, and C. Wilson (1998). *When is Less More? Faithfulness and Minimal Links in wh-Chains*, pp. 249–289. MIT Press.
- Partee, B. H. (1998). Copula inversion puzzles in English and Russian. In *Formal Approaches to Slavic Linguistics: The Seattle Meeting*, Ann Arbor, MI, pp. 361–395. Michigan Slavic Publications.
- Pereltsvaig, A. (2001). *On the Nature of Intra-Clausal Relations: A Study of Copular Sentences in Russian and Italian*. Ph. D. thesis, McGill University.
- Richards, N. (2006). A distinctness condition on linearization.
- Walter, M. A. (2007). *Repetition Avoidance in Human Language*. Ph. D. thesis, Massachusetts Institute of Technology.
- Yip, M. (1998). *Identity Avoidance in Phonology and Morphology*, pp. 216–246. Stanford, CA: CSLI Publications.

A A Minimalist analysis

- Tools
 - PF constraint against consecutive homophonous *wh*-words DPs
 - copy theory of movement - head-deletion preference (Franks 1998)
 - Bošković's data
- (12) a. $N=\{\check{t}o_2, obuslovilo_1, V_1, v_1 T_1, C_1\}$
- | | |
|--|---|
| [_{VP} obuslovilo $\check{t}o_j$] | merge <i>V</i> , <i>obuslovilo</i> , and $\check{t}o$ |
| [_{VP} $\check{t}o_i$ $\check{t}o_j$ [_{V'} $\check{t}o_i$ obuslovilo [_{VP} obuslovilo $\check{t}o_j$]]] | merge <i>v</i> and $\check{t}o$ |
| spellout: [_{VP} $\check{t}o_i$ $\check{e}t\check{o}_j$ [_{V'} $\check{e}t\check{o}_j$ obuslovilo [_{VP} obuslovilo $\check{t}o_j$]]] | |
| [_{TP} [_{VP} $\check{t}o_i$ $\check{e}t\check{o}_j$ [_{V'} $\check{e}t\check{o}_j$ obuslovilo VP]]] | merge <i>T</i> |
| [_{TP} $\check{t}o_i$ [_{VP} $\check{t}o_i$ $\check{e}t\check{o}_j$ [_{V'} $\check{e}t\check{o}_j$ obuslovilo VP]]] | check <i>T</i> 's ϕ , case agreement |
| [_{CP} [_{TP} $\check{e}t\check{o}_j$ [_{VP} $\check{t}o_i$ $\check{e}t\check{o}_j$ [_{V'} $\check{e}t\check{o}_j$ obuslovilo VP]]] | merge <i>C</i> |
| [_{CP} $\check{t}o_i$ $\check{e}t\check{o}_j$ [_{TP} $\check{e}t\check{o}_j$ [_{VP} $\check{t}o_i$ $\check{e}t\check{o}_j$ [_{V'} $\check{e}t\check{o}_j$ obuslovilo VP]]] | check <i>C</i> 's wh |
| spellout: [_{CP} $\check{t}o_i$ $\check{e}t\check{o}_j$ [_{TP} $\check{e}t\check{o}_j$ [_{VP} $\check{t}o_i$ $\check{e}t\check{o}_j$ [_{V'} $\check{e}t\check{o}_j$ obuslovilo VP]]]]] | |
- b. converges: [_{CP} $\check{t}o_i$ $\check{e}t\check{o}_j$ [_{TP} $\check{e}t\check{o}_j$ [_{VP} $\check{e}t\check{o}_j$ [_{V'} $\check{e}t\check{o}_j$ obuslovilo [_{VP} obuslovilo $\check{t}o_j$]]]]]
- Obligatory *est'*⁷
 - Derivation without *est'* in (13) crashes, derivation with *est'* in (14) converges
- (13) a. $N=\{kto_2, T_1, C_1\}$
- | | |
|---|---|
| [_{DP} kto_i kto_j] | merge <i>kto</i> and <i>kto</i> |
| [_{TP} [_{DP} kto_i kto_j]] | merge <i>T</i> |
| [_{TP} kto_i [_{DP} kto_i kto_j]] | check <i>T</i> 's ϕ , case agreement |
| [_{CP} [_{TP} kto_i [_{DP} kto_i kto_j]]] | merge <i>C</i> |
| [_{CP} kto_i kto_j [_{TP} kto_i [_{DP} kto_i kto_j]]] | check <i>C</i> 's wh |
- b. CRASHES: [_{CP} kto_i kto_j [_{TP} kto_i [_{DP} kto_i kto_j]]]
- (14) a. $N=\{kto_2, est'_1, T_1, C_1\}$
- | | |
|--|---|
| [_{DP} kto_i kto_j] | merge <i>kto</i> and <i>kto</i> |
| [_{TP} est' [_{DP} kto_i kto_j]] | merge <i>T</i> and <i>est'</i> |
| [_{TP} kto_i est' [_{DP} kto_i kto_j]] | check <i>T</i> 's ϕ , case agreement |
| [_{CP} [_{TP} kto_i est' [_{DP} kto_i kto_j]]] | merge <i>C</i> |
| [_{CP} kto_i kto_j [_{TP} kto_i est' [_{DP} kto_i kto_j]]] | check <i>C</i> 's wh |
- b. converges: [_{CP} kto_i kto_j [_{TP} kto_i est' [_{DP} kto_i kto_j]]]
- With phases

⁷I utilize the structure provided by Pereltsvaig (2001).



- Again, derivation without *est'* in (15) crashes, derivation with *est'* in (16) converges

- (15) a. N={bogatyj₂, vrač₂, T₁, C₁}
[_{DP} [_{DP} bogatyj vrač]_i [_{DP} bogatyj vrač]_j] merge *bogatyj vrač*
[_{TP} [_{DP} [_{DP} bogatyj vrač]_i [_{DP} bogatyj vrač]_j]] merge T
[_{TP} [_{DP} bogatyj vrač]_i [_{DP} [_{DP} bogatyj vrač]_i [_{DP} bogatyj vrač]_j]] check T's ϕ , case agreement
b. CRASHES: [_{TP} [_{DP} bogatyj vrač]_i [_{DP} ~~[_{DP} bogatyj vrač]_i~~ [_{DP} bogatyj vrač]_j]]
- (16) a. N={bogatyj₂, vrač₂, est' T₁, C₁}
[_{DP} [_{DP} bogatyj vrač]_i [_{DP} bogatyj vrač]_j] merge *bogatyj vrač*
[_{TP} est' [_{DP} [_{DP} bogatyj vrač]_i [_{DP} bogatyj vrač]_j]] merge T and est'
[_{TP} [_{DP} bogatyj vrač]_i est' [_{DP} [_{DP} bogatyj vrač]_i [_{DP} bogatyj vrač]_j]] check T's ϕ , case agreement
b. converges: [_{TP} [_{DP} bogatyj vrač]_i est' [_{DP} ~~[_{DP} bogatyj vrač]_i~~ [_{DP} bogatyj vrač]_j]]

- Advantages

- Under this analysis, a single constraint can handle all the data
- This doesn't rely on look-ahead, and it can look back only as far as SS (I think) (cite P and Z)

- Problems

- This constraint seems stipulative, you might imagine that it should be derived
- This constraint would need to apply at a level where the relevant phonological, syntactic, and linearity information would be available, and it's not obvious that such a level exists

B A distinctness analysis

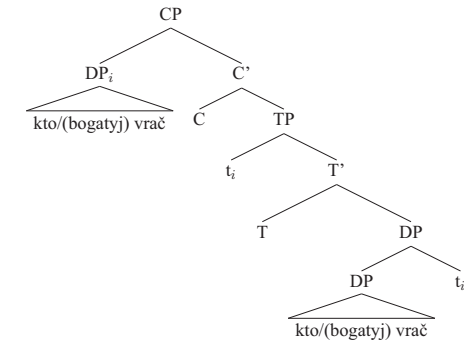
- Tools

- Linear Correspondence Axiom: $d(A)$ is a linear ordering on T (i.e. there is a correlation between hierarchical structure and linear order) (Kayne 1994)
- Distinctness: If a linearized statement $\langle \alpha, \alpha \rangle$ is generated, the derivation crashes. (Richards 2006)
 - o α can be assumed to refer to syntactic feature bundles.
 - o phases: CP, transitive vP, PP, KP

- So in Russian, we might have something like...

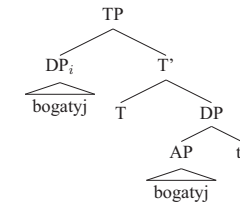
- In (17) there is the pair $\langle D, D \rangle$, which are presumably too featurally similar.

- (17)



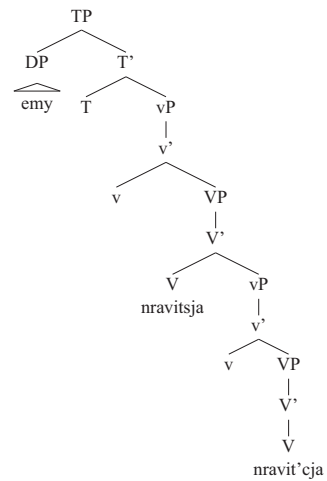
- In (18), there is no $\langle \alpha, \alpha \rangle$

- (18)



- In (19), presumably $\langle V, V \rangle$ and $\langle v, v \rangle$ involve sufficient featural distinctions (note the higher verb is finite, the lower verb is not).

(19)



- Additional supporting evidence for this type of analysis
 - Typically, Russian equative sentences, which presumably involve featurally similar DPs, require *éto*, which presumably has a different structure from (17).
 - e.g. Mark Twain - *(éto) Samuel Clemens (Geist 2008)
- Predictions?
 - Homophony is irrelevant in the data above, it's featural similarity and syntax that matters. So, non-homophonous copular sentences like (4a) should be ungrammatical if they are featurally similar enough. If homophonous sentences like (4a) can be interpreted in a less featurally similar way, they should be grammatical.
- Problems with this analysis
 - It's not clear why null vs. overt copula should make a difference (they presumably use the same structure, Pereltsvaig 2001)
 - It's not clear that a level exists where the relevant syntactic and linearity information would be available.