Remarks and Replies

On the Difference between the Lexicon and Computation (Regarding Slavic Yers)

Tobias Scheer

This article discusses three concerns regarding Becker and Gouskova's (2016) analysis of Russian yers that relies on cluster-based yer vocalization and two sublexicons (morphemes with and without yers), to which lexically specific constraints refer. First, it misses the basic generalization about Slavic yers expressed by the established analysis (Lower): yer vocalization is triggered by five different mechanisms, instead of one mechanism under Lower. It is further shown that the major objection against Lower disappears when the existence of final empty nuclei is recognized. Second, Becker and Gouskova confound generalizations about the lexical distribution of yers in morphemes and the computational mechanism that decides which yers appear on the surface. They argue that Lower was established before relevant cluster-based generalizations were discovered, hence misses out on relevant empirical material that invalidates its central idea, that clusters are irrelevant for yer vocalization. However, the phenomena their argument is based on do not concern yer vocalization (computation): they are lexical in kind and therefore confirm the irrelevance of clusters for yer vocalization, supporting Lower. Third, although generalizations about yer-deletion-created clusters are central for Becker and Gouskova's analysis, they are irrelevant for learners (children or adults). The authors' experimental evidence precisely shows that speakers are happy to lexicalize and compute sequences (such as yerCC) that are absent from the lexicon. The gaps at hand are thus accidental, rather than systematic.

Keywords: Slavic yers, vowel-zero alternations, lexicon vs. computation, Government Phonology, final empty nuclei, morpheme structure constraints.

1 Lexical Patterns Are Part of the Knowledge of Russian Speakers

Becker and Gouskova (2016) (following Gouskova and Becker 2013) show that generalizations that can be made over the lexicon regarding vowels that alternate with zero in Russian (so-called

Many thanks to Gerhild Zybatow for proofreading the Russian data, and to two anonymous reviewers for valuable comments that have greatly improved the article.

Linguistic Inquiry, Volume 50, Number 1, Winter 2019
197–218

doi: 10.1162/ling_a_00303

yers) are part of grammatical knowledge: speakers reproduce them productively when rating the grammaticality of nonce words (in a wug test). For example, vowels that precede a root-final consonant cluster are never yers (i.e., never alternate with zero). Clusters are common word-finally (-st#: *pomóst* 'stage Nsg') and three-membered clusters may be created by the deletion of a yer (stoC# – stC-V: *kost'ór – kostr-á* 'fire Nsg, Gsg').[1] But in no Russian word is the vowel in a -CoCC# root deleted (**pomóst – pomst-á*). This lexical distribution, which the authors call complex coda blocking (there are no yers before root-final clusters), is mirrored by the fact that in nonce words, speakers accept deleted vowels significantly more often when they precede root-final singleton consonants (nonce *pišoch – pišch-a*) than when they are followed by a root-final cluster (nonce *pišochl – pišchl-a*). In order to make these judgments, Becker and Gouskova (2016) argue, speakers necessarily call on their grammatical knowledge.

Becker and Gouskova (2016) also show that the absence of yers before root-final clusters is a statement about the source (input) that cannot be captured by any product-based (output-based) generalization: -CCC-V strings are fine if they come from a -CCoC# input, but not if they come from a -CoCC# input. The authors further demonstrate that (unlike other source-oriented generalizations), this pattern cannot be expressed in terms of faithfulness. Rather, it needs to be encoded by a mechanism akin to *SPE*'s morpheme structure constraints (Chomsky and Halle 1968) that defines the properties of lexical items by running candidates for lexicalization through a specific set of constraints (which the authors call the gatekeeper grammar) that make sure that no lexical entries with certain properties occur (in the present case, -CyerCC#). That is, speakers have two grammars, a gatekeeper grammar (GK) and a grammar proper (GP), the latter mapping inputs onto outputs.

Since the constraint against roots with final clusters is specific to those roots that have a yer (final clusters happily occur in non-yer roots), Becker and Gouskova (2016:407) devise two separate sublexicons, one containing all and only those morphemes that exhibit a yer, the other containing morphemes where no yers occur. Speakers thus have two gatekeeper grammars, GK_{yer} and $GK_{non\text{-}yer}$. The constraint ranking of the former rules out roots with final clusters, while the grammar of the latter does not.

Another piece of speakers' knowledge that Becker and Gouskova's (2016) nonce word test reveals is that yer deletion is preferred if the resulting cluster is (C)TR, as compared to (C)RT, (C)TT, (C)RR (where *T* represents obstruents and *R* sonorants). Hence, speakers prefer the vowel to be deleted in (nonce) *sóm* (yielding *sm-á*) when compared to (nonce) *mós* (yielding *ms-á*). Becker and Gouskova (2016:394) report that this generalization is not mirrored in the lexicon. It is also different in kind with respect to the dispreference for yers followed by root-final clusters

[1] In this article, Russian words are rendered in a transliteration that does not indicate vowel reduction (i.e., that is oblivious to the issues discussed). The secondary palatal articulation of consonants is marked by an apostrophe when palatality is not predictable from the following vowel (*s*' is [sʲ]), including before *ë* (= jo) as in *осёл* 'donkey' (transcribed as *os'ól*). Stress is indicated by an acute accent on the vowel, and *š*, *ž*, *c* represent [ʃ], [ʒ], [ʦ], respectively.

since it has a clear output-driven (product-oriented in the authors' terms) rationale based on sonority.

Here again, the authors build on the split of the Russian lexicon into one sublexicon that contains yer-bearing morphemes and another sublexicon that contains morphemes without yers. Since the generalization is product-oriented and hence falls into the competence of grammar proper, there are two lexically specific versions, GP_{yer} and $GP_{non\text{-}yer}$.

A Russian speaker thus operates with two sublexicons and four grammars in total (for masculine nouns): GK_{yer}, $GK_{non\text{-}yer}$, GP_{yer}, and $GP_{non\text{-}yer}$. As a result, Becker and Gouskova (2016) attribute the absence of roots that end in yerCC# and the preference for yer-deletion-created clusters of rising sonority to entirely different mechanisms. While both generalizations are part of speakers' grammatical knowledge, the locus of this knowledge is the gatekeeper grammar in the former case, but the grammar proper in the latter. Finally, note that the two sublexicons are also motivated by the fact that in Russian (as in other Slavic languages), the occurrence of yers in lexical items is unpredictable: *e* and *o* may (*rót – rt-á* 'mouth Nsg, Gsg') or may not (*pót – pót-a* 'sweat Nsg, Gsg') alternate with zero. Following Gouskova (2012), Becker and Gouskova (2016) encode this contrast by lexically marking classes of morphemes (yer morphemes vs. non-yer morphemes), rather than individual segments (regular vowels vs. yers).

2 Three Points

In this reply, I would like to make three points.

1. *The basic generalization regarding Slavic yers is missed.* Gouskova's (2012) and Becker and Gouskova's (2016) way of approaching Slavic vowel-zero alternations in general and their Russian incarnation in particular misses a simple and well-established generalization that covers all cases of the intricate pattern at stake: yers appear on the surface if and only if the following vowel is also a yer. This is what a massive body of literature (on which more in section 4) has established as the so-called Lower rule since Lightner 1965. Instead of holding the avoidance of certain clusters responsible for the appearance of vowels on the surface (according to Becker and Gouskova, vowels are interpreted as non-yers in RoT-V, ToT-V, and RoR-V in order to avoid the clusters that would result from deletion), Lower builds on a lateral (syntagmatic) relation with the following nucleus. Clusters play no role, and there is no need for sublexicons or specific grammars that enforce lexical patterns.

2. *Phenomena that concern the lexical distribution of yers in morphemes are misattributed to the computational system (i.e., the mechanism that decides which yers appear on the surface).* This is relevant on a number of occasions in Gouskova 2012, Gouskova and Becker 2013, and Becker and Gouskova 2016 (to be discussed in section 5), especially since the authors make arguments against Lower and in favor of the cluster-based analysis that are based on what they incorrectly take to be computational.

For example, the fact that TR clusters are preferred over other clusters when resulting from yer deletion is a fact about how ambiguous vowels (that could be either a yer or a regular stable vowel) are lexicalized (in new loans, acronyms, nonce words, etc.). It has nothing to do with the computational mechanism (Lower in the established literature), which decides about the surfacing

of yers once they are recorded in the lexicon. The workings of this computation are 100% regular no matter the sonority of the cluster created. It is this computational mechanism that is specific to Russian (or other Slavic languages) and that needs to be encoded in its grammar—crucially, without any reference to the sonority of resulting clusters. In other words, the grammar proper, where input-output mapping is carried out, must not make reference to the sonority of resulting clusters since this factor is irrelevant.

3. *Generalizations about the sonority profile of yer-deletion-created clusters are irrelevant for human learners (children or adults, as opposed to the machine learners that are used in Becker and Gouskova 2016).* When faced with a word that is unknown and whose last vowel is *e* or *o*, learners of Russian need to make a decision about how the surface *e*,*o* is stored in the lexicon: either as stable *e*,*o* or as alternating *e*,*o* (i.e., a yer). Likewise, when root-final clusters are encountered in new words that learners have only come across in a suffixed form [. . . CC-V], they need to be interpreted as either bearing a yer (/ . . . CyerC/) or not (/ . . . CC/) upon lexicalization. These decisions need to be made by children in (first) language acquisition, and by adults when confronted with loans, acronyms, and so on. Children and adults may get it wrong; Łukaszewicz (2006:15–16) reports the child data in table 1 from Polish (where the workings of yers are the same as in Russian).

All examples involve final TR# clusters (which, according to Becker and Gouskova's findings, have the highest likelihood of inducing the lexicalization of a yer in their midst), which the child obviously cannot produce: as table 1 shows, they are either broken up by a yer (a) or simplified (b). That is, the child lexicalizes either a yer (/wiatъr/) or a cluster (/Piotr/), whose second consonant is lost when occurring in word-final position. Łukaszewicz reports that the choice of the two strategies is unpredictable because, she argues, the child knows that final (TR-) clusters may (*sweter – swetr-a* 'sweater Nsg, Gsg') or may not (*filtr – filtr-a* 'filter Nsg, Gsg') host a yer: there is nothing in the available data that a decision could be based on, and what ends up being lexicalized is a matter of chance.

Sooner or later, however, children (or adults) will be exposed to relevant evidence that disambiguates the situation. When hearing Russian *bobr-á* 'beaver fur Gsg' and *dolg-á* 'long

Table 1
Polish acquisition (A., 3;11) (Łukaszewicz 2006:15–16)

	Adult target		Child production		
	Nsg	Gsg	Nsg	Gsg	
a. TR#: Yer insertion	wiatr	wiatr-u	wiater	same as adult	'wind'
	motocykl	motocykl-a	motocykel	same as adult	'motorcycle'
	bóbr	bobr-a	bober	same as adult	'beaver'
b. TR#: Deletion of final consonant	pomysł	pomysł-u	pomys	same as adult	'idea'
	zgadł (masc.)	zgadł-a (fem.)	zgad	same as adult	'he/she guessed'
	Piotr	Piotr-a	Piot	same as adult	'Peter'

fem.', children have no way to know whether or not there is a yer hidden in the root-final cluster.[2] They may make a guess based on the knowledge that Becker and Gouskova have shown natives possess: they will then be more likely to suppose that the root meaning 'beaver fur' has a yer (since the cluster created by yer deletion is of the preferred TR type) and hence its Nsg is *bob'ór*. By contrast, the root-final RT cluster in *dolg-á* drives children into choosing an underlying form without a yer, which will produce the masculine form *dólg*. Children will have been right in the former case, but wrong in the latter: the masculine form of *dolg-á* is *dólog*. When children come across this masculine form, though, they understand that having lexicalized /dolg-/ was a mistake and correct the lexical entry to /dolъg-/. The same goes for adults.[3]

In this entire process, whether or not a preference for yer-deletion-created TR clusters is part of speakers' knowledge is entirely irrelevant: it does not help learners in any way, nor does it alter the learning path. If anything, the bias for yer deletion in /(C)TyerR/ (as compared to /(C)RyerT/, /(C)TyerT/, /(C)RyerR/) that Becker and Gouskova have identified in their experiment will induce children (and adults) into error more often than if they went by chance when they make guesses whether the root of a surface √-TR-V# sequence ends in /-TR/ or in /-TyerR/. This is because, as Becker and Gouskova (2016:394) report, the preference observed in the experiment is not mirrored in the actual lexicon of Russian: vowel deletion is not particularly probable in (C)To/eR# items. Hence, putting the experimentally established knowledge in favor of a /(C)TyerR-V/ – [(C)TR-V] mapping to use is dysfunctional: learners will tend to interpret [(C)TR-V] items as yer-bearing although this mapping is rather infrequent in the actual lexicon.

It would be interesting to see whether the TR-favoring bias that Becker and Gouskova have found when asking adults to judge unknown words is also present in the actual behavior of children and adults when they lexicalize new words in the wild.

In sum, the bias in favor of yer-deletion-created TR clusters is alien to the workings of Russian: it plays no role in either the computational system or the lexicon and is irrelevant in the acquisition of new words (by human learners, both children and adults). If anything, it is functionally counterproductive in the latter area.

What is the origin and locus of the sonority-based bias in Russian speakers, then? Showing that sonority sequencing is ubiquitous in productive phonological processes, that it is supported by typological data, and that it constrains the behavior of speakers in psycholinguistic experiments, Berent (2013:165–198) argues that sonority sequencing is a grammatical universal since it cannot be derived from extragrammatical factors (such as phonetics). Berent also shows that sonority sequencing does not merely extend to lexical items that speakers have never come across: it is also active in structures that are unattested in the speaker's language (e.g., branching onset preferences

[2] A reviewer points out that there is documented evidence from Korean showing that learners can override or ignore data in their native environment and develop coexisting morphophonemic variants (Ito 2010, Do, Ito, and Kenstowicz 2014). As far as I can see from the literature and my own experience, this is not the case in Slavic regarding yer vocalization. That is, no one will say or accept Polish **swetr* 'sweater Nsg' or **filter* 'filter Nsg' (only *sweter* and *filtr* occur and are judged well-formed).

[3] Here and below, I use the traditional symbols ъ and ь to refer to yers in lexical representations in modern languages, as well as to regular Common Slavic vowels in Common Slavic words.

produced by Korean speakers, whose language lacks branching onsets; Berent et al. 2008). Given this backdrop, the behavior of Russian speakers in Becker and Gouskova's (2016) experimental data may be a kind of emergence of the unmarked: the preference for TR clusters is irrelevant in Russian (grammar, lexicon, and learning), but still part of the universal knowledge of Russian speakers.

Optimality Theory aims to encode all knowledge that speakers have of their language in a single locus, the constraint hierarchy (Prince and Smolensky 2004 [1993]:chap. 9). This constraint hierarchy is then responsible for actual phonological computation (input-output mapping) as much as for static properties such as inventory definition, systemic properties of inventories, parameter settings, or even the regularities that are found in the lexicon (Richness of the Base: grammatical inputs are universal). It also encodes nonphonological processes (i.e., those that do not relate a phonological input to a phonological output) such as analogy (output-output faithfulness, paradigm uniformity) or the workings of the interface(s). In alternative approaches, active phonological computation (i.e., where phonological inputs are mapped onto phonological outputs) is not mixed with the other types of knowledge. In the theory of contrast promoted by Dresher (2009) and Hall (2011), for example, systemic properties are acquired and then stored as stable information in the form of a contrastive hierarchy to which phonological computation makes reference—but which is not a piece of phonological computation itself. The same goes for parameters and inventory definition, which may be hardwired knowledge that is stored independently of phonological computation. In other words, the computational mechanism that maps inputs onto outputs may not be all there is regarding grammar, and it may not be the only locus where the knowledge that speakers have about their language is encoded.

This is how the above conclusion is to be understood: the mechanism that carries out input-output computation in Russian must not include any bias in favor of yer-deletion-created TR clusters. The pattern thus provides evidence for the distribution of tasks over different loci in grammar (rather than their exclusive concentration in input-output computation), just as Becker and Gouskova's (2016) central distinction between source-oriented and product-oriented generalizations (and hence between a gatekeeper grammar and a grammar proper) does.

Points 1 and 2 above are further contextualized below.

3 Cluster-Based Analyses

A question that has been debated at length in the literature on Slavic yers is whether alternating vowels are underlyingly absent and inserted, or present and deleted. Insertion-based analyses where certain difficult clusters are held to trigger epenthesis have been proposed by, among others, Laskowski (1975, Polish), Czaykowska-Higgins (1988, Polish), Piotrowski (1992, Polish), and Townsend (1975:62–65, Russian).

These analyses have been refuted because there is no context for insertion: in Polish, for example, *-tr#* in /futr/ 'fur Gpl', *-p^{j}#* in /stop^{j}n/ 'step Nsg', and *-mn#* in /trum^{j}n/ 'coffin Gpl' are broken up and the words appear as *futer*, *stop^{j}en*, and *trum^{j}en*, respectively. However, the same word-final clusters do happily exist in Polish without being broken up: *wiatr* 'wind Nsg', *wapń* 'calcium Nsg', *hymn* 'hymn Nsg'. This pattern is pervasive in Slavic and led to the abandon-

ment of the insertion analysis. Relevant literature includes Pesetsky 1979 (Russian), Gussmann 1980:26–28 (Polish), Rubach 1984:28–29, 1993:134–144, 2013:1141 (Polish and Slovak), Farina 1991:256–257 (Russian), Szpyra 1992:280–292, 1995:94–99 (Polish), and Yearley 1995:538 (Russian). For overviews, see Bethin 1998:210–211 and Scheer 2011.

Starting with Szpyra 1992, there is a line of thinking that records yers in lexical representations (and hence abandons insertion) but, as in the insertion perspective, holds that the vocalization of yers (i.e., their promotion to surface existence) is due to the avoidance of certain consonant clusters. Yearley 1995 is an incarnation of this idea, but the author does not explain how the yer vocalization mechanism actually distinguishes between identical clusters that are broken up in some words but not in others. It is only Gouskova (2012) who makes the cluster-based analysis technically workable by devising morpheme-specific constraints (e.g., Pater 2009) that apply to two separate classes of morphemes (or sublexicons), one with words that contain yers, the other with words that do not. Thus, the Russian root meaning *weasel*, /lasъk$_{yer}$/, belongs to the yer lexicon, while the root meaning *caress*, /lask$_{non\text{-}yer}$/, does not. Constraints that apply only to the yer lexicon then rule out word-final *-sk#* clusters, to the effect that the yer in /lasъk/ 'weasel Gpl' must appear on the surface: *lások* (Nsg *lásk-a*). Final *-sk#* clusters in the non-yer lexicon are not affected since the constraints against them do not control this set of morphemes. Therefore, the Gpl of /lask$_{non\text{-}yer}$/ 'caress' appears as *lásk*. The Nsg of both words is *lásk-a*, though.

This is where the idea of a split in the Russian lexicon—namely, into yer- and non-yer sublexicons—comes from. Note that what is lexically marked in Gouskova's analysis is not the fact that a given vowel is or is not a yer (which is what regular analyses do), but the fact that a given morpheme does or does not bear a yer—hence the title of Gouskova 2012, "Unexceptional Segments," referring to the fact that all [ɛ]'s and all [ɔ]'s that appear on the surface in Russian have the same lexical identity regardless of whether or not they alternate with zero. For example, the lexical entry of the aforementioned *rót – rt-á* 'mouth Nsg, Gsg' will be /rot$_{yer}$/, while that of *pót – pót-a* 'sweat Nsg, Gsg' will be /pot$_{non\text{-}yer}$/ (instead of /rъt/ vs. /pot/ as on the segment-based analysis).[4]

Like insertion-based approaches, Yearley's (1995) and Gouskova's (2012:83) analysis is cluster-based: yers appear on the surface because certain clusters are illegal. On their analysis, the following clusters trigger yer vocalization in Russian. A yer appears on the surface when it is the only vowel in the word (/sъn/ → *són* 'dream Nsg' instead of **sn*), since there are no words without vowels in the language. A yer is also vocalized when occurring in a word-final cluster that may (/lasъk/ → *lások* 'weasel Gpl') or may not be attested in the language.[5] Gouskova (2012:83) calls the latter "unpronounceable": the yers in /chlopъk/ and /korotъk/ surface in

[4] See Rubach 2013 for a critique of the analysis based on the lexical marking of yer morphemes, showing that this approach runs into insuperable problems when applied to Polish. A reviewer points out that the split into a yer and a non-yer lexicon may (or ought to) have consequences elsewhere in the phonology of Russian (regarding stress for example), since it is plausible that such an overarching lexical distinction will be exploited by computation.

[5] For the sake of exposition, I continue using yers in underlying forms even when discussing Gouskova 2012, which has no yers.

chlópok ‘cotton Nsg’ (cf. *chlópk-a* ‘cotton Gsg’) and *kórotok* ‘short masc.’ (cf. *korotk-á* ‘short fem.’) because their absence would create *-pk#* and *-tk#*, clusters that are unattested in Russian.

Yearley (1995:545, echoed in Gouskova 2012:87–88) further wonders why the two word edges seem to produce opposite behavior in Russian yers. Yers systematically surface to break up word-final clusters (/lasъk/ → *lások* ‘weasel Gpl’), but are never realized when word-initial clusters could be avoided (/l’ьst-it’/ → *l’st-ít’* ‘to flatter’; cf. *lést’* ‘flattery’). On Yearley’s (and Gouskova’s) analysis, this contrast stems from the ability of initial clusters to have their edgemost consonant adjoined to the prosodic word (i.e., bypassing the onset and the syllable node), while this option is denied to word-final clusters.

Finally, there is another reason why yers vocalize: Triconsonantal Cluster Blocking (Gouskova 2012:89–91) prevents yers from deleting (which, on Yearley’s and Gouskova’s basic analysis, they should) when their absence would create three consonants in a row. Both word-initial and word-medial three-member clusters are subject to this restriction since, Gouskova argues, they are unsyllabifiable (note the difference with “unpronounceable” clusters mentioned earlier). For example, the last vowel of *kot’ól* ‘boiler Nsg’ is a yer because it is absent in Gsg *kotl-á.* The diminutive of this word is *kotel-ók* in Nsg and *kotel-k-á* in Gsg. The reason why the last root vowel in the latter form does not delete, Gouskova contends, is that deletion would produce the illegal three-member cluster *-tlk-*.[6] This analysis applies to all sequences of two or more yers, which are frequent and characteristic for double diminutives (e.g., *igól-k-a – igól-ok* ‘needle dim. Nsg, Gpl’, *igól-oč-k-a – igól-oč-ek* ‘needle double dim. Nsg, Gpl’).

4 Lower

4.1 A Unified Analysis

Yers in word-initial and word-final clusters show contrastive behavior only under the assumption that yer vocalization is cluster-driven. The issue disappears under the established analysis of yers that originated in Lightner 1965 and has had numerous offspring in linear and autosegmental incarnations. This analysis is known as Lower, and its basic insight is that yers vocalize if and only if they are followed by another yer in the underlying representation (Gussmann 1980, Rubach 1984, 1986, Kenstowicz and Rubach 1987, Farina 1991, and many others; see the survey in Scheer 2011).

Lower offers a unified perspective where all yers are vocalized for the same reason and are governed by the same mechanism. When the patterns are viewed through this lens, there is no such thing as opposite behavior of yers at the two word edges, and no specific edge-based mechanism (the ability of edgemost consonants to adjoin to the prosodic word) needs to be devised.

[6] On Gouskova’s analysis, there are also syllabifiable medial triconsonantal clusters, which can be happily created by yer deletion. Examples are the aforementioned *kost’ór – kostr-á* ‘fire Nsg, Gsg’ and *dólž-en – dolž-n-á* ‘must pred. adj., masc., fem.’. It is not clear what the rationale is that distinguishes syllabifiable from unsyllabifiable clusters; Gouskova does not discuss this issue.

Yers after word-initial consonants are followed by a stable vowel and thus remain unvocalized (/mъx-a/ → *mx-á* ‘moss Gsg’), while yers before word-final consonants are followed by another yer and therefore appear on the surface (/lasъk-ъ/ → *lások* ‘weasel Gpl’). The same goes for yer chains: the leftmost yer is followed by another yer and therefore vocalizes, while the rightmost yer deletes in case it is followed by a stable vowel (/kotъl-ъk-a/ → *kotel-k-á* ‘boiler dim. (= kettle) Gsg’). Clusters (or syllabification) play no role in this analysis.

Hence, the amorphous set of reasons that is held responsible for yer vocalization under the cluster-based analysis reduces to a single unified reason and a single mechanism when Lower is assumed.[7] A sample of relevant contexts comparing the cluster-based analysis and Lower is given in table 2.

Abandoning Lower for the reasons discussed by Yearley (1995:536–537) and Gouskova (2012:109–110) thus comes at the cost of missing an overarching generalization: we are facing a single phenomenon, yers and their vocalization, rather than a series of unrelated processes. Also recall that dividing the lexicon into morphemes that do vs. do not contain yers is as unnecessary under the workings of Lower as the assumption of sublexicon-specific computational instructions. Finally, note that Gouskova (2012:110) points out that it is impossible to implement Lower in Optimality Theory since Richness of the Base does not allow for a situation where all lexical items of a language end either in a full vowel or in a yer.

The major objection that is leveled against Lower, regarding the existence of word-final yers that never appear on the surface, is addressed in the remainder of this section.

Table 2
Vowel-zero alternations: Amorphous set of reasons (cluster-based) vs. one single reason (Lower)

Yers are		Reason for yer (non)vocalization	
Vocalized	Nonvocalized	Cluster-based	Lower
a. /sъn-ъ/ → *són*	/sъn-a/ → *sn-á*	No vowelless words	yerC_0yer
b. /lasъk-ъ/ → *lások*	/lasъk-a/ → *lásk-a*	No yer-deletion-created final cluster	yerC_0yer
c. /chkopъk-ъ/ → *chlópok*	/chlopъk-u/ → *chlópk-u*	Unpronounceable final cluster	yerC_0yer
d. /mъx/ → *móx*	/mъx-a/ → *mx-á*	#C adjoins to the prosodic word	yerC_0yer
e. /kotъl-ъk-a/ → *kotel-k-á*	/kotъl-a/ → *kotl-á*	Triconsonantal Cluster Blocking	yerC_0yer

[7] Note that the basic generalization expressed by Lower may be altered by cyclic effects produced by prefixes. This is orthogonal to the scope of Lower and also challenges alternative yer-vocalizing mechanisms (see Pesetsky 1979:sec. 1.2, Farina 1991:365–382, Yearley 1995:557–569).

4.2 Word-Final Yers That Never Appear on the Surface

The unification-based argument in favor of Lower hinges on the presence of yers that follow word-final consonants: in /sъn-ъ/ → *són*, for example, vocalization occurs because the root /sъn/ is followed by a yer. In the classical literature, yers that occur after word-final consonants and never appear on the surface (because they are not followed by another yer) are interpreted as case markers.

One motivation of insertion-based analyses was to get rid of these "abstract" word-final vowels (e.g., Szpyra 1992:302–303); these are also taken as the central argument by Yearley (1995:536–537) and Gouskova (2012:108n25) for abandoning Lower. This objection disappears in the environment of Government Phonology, where vowels that alternate with zero are lexically distinct from the items that occur after word-final consonants: the former are floating pieces of melody, while the latter are (final) empty nuclei that come for free since word-final consonants are onsets of empty nuclei in all languages. To make this clear, the following section introduces autosegmental implementations of Lower and yers.

4.3 Yers as Nuclei with a Floating Piece of Melody

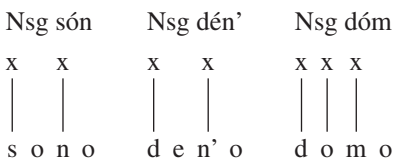
The fundamental insight of Lower is that vowel-zero alternations in Slavic are the result of a regressive (right-to-left) intervocalic relation. The target is the leftmost vowel, whose phonetic value is determined by the neighbor to its right: the former appears on the surface if the latter is a yer; otherwise, it remains unexpressed. In Government Phonology (Kaye, Lowenstamm, and Vergnaud 1990, Lowenstamm 1996, Scheer 2004), this kind of head-final lateral relation between two nuclei is called government: the target nucleus is phonetically absent when governed (i.e., when the lateral relation is established), but appears on the surface in case it escapes government. These workings are illustrated below.

The first autosegmental implementation of Lower is due to Hyman (1985:58–59) and Rubach (1986). On their analysis, yers are floating pieces of melody (segments) that are devoid of syllabic support in the lexicon (while stable vowels are lexically associated to an x-slot or a mora). In the government-based approach, Scheer (2004:§§81–82, 2005) and Gussmann (2007) follow the idea that alternating vowels (yers) are floating, while stable vowels are lexically associated. Unlike on Hyman's and Rubach's approach, the nucleus of the former items is thus present.

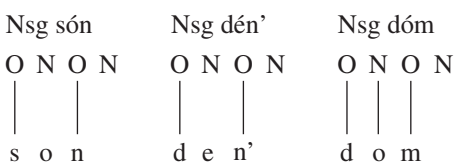
The original and government-based representations of yers and stable vowels are shown in (1), using Russian *són – sn-á* 'dream Nsg, Gsg' for alternating vowels and *dóm – dóm-a* 'house Nsg, Gsg' for nonalternating ones. On the autosegmental account, specific yer vowels (ъ, ь) that are part of the phonemic inventory of the language do not need to be assumed. As shown in (1a–b), yers are segmentally indistinguishable from stable vowels: alternating [ɔ] (back yer) in *són – sn-á* is segmentally /o/ just as much as nonalternating [ɔ] in *dóm – dóm-a.* The difference lies in the autosegmental relationship yers and non-yers enjoy with syllabic positions (for an overview, see Bethin 1998:206–214): on the classical approach, (1a), yers are floating segments that do not possess any syllabic position (x-slot), while stable vowels are lexically associated to a syllabic position.

(1) *Underlying representation of yers and stable vowels*

a. Hyman 1985, Rubach 1986



b. Scheer 2004, 2005



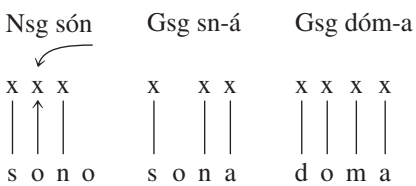
The government-based analysis under (1b) follows the same logic, except that floating yers do possess a nucleus in the lexicon: the only thing that differentiates them from stable vowels is the absence of association to their nucleus.

Finally, recall from section 4.2 that on the original approach, word-final consonants are followed by yers, which represent case markers (hence the floating *o* after word-final consonants in (1a)). In the government-based alternative, these final yers do not exist: instead, there is a final empty nucleus, which is crucially distinct from a yer since it does not possess any floating segment.[8] This is the central point to be illustrated here since final yers are the major reason why cluster-based approaches reject Lower.

The way yer vocalization works, given these lexical representations, is shown in (2).

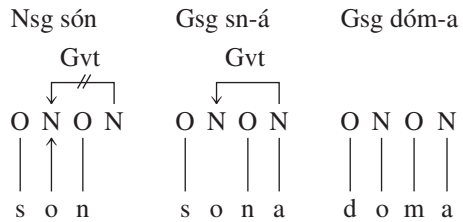
(2) *Yer vocalization*

a. Hyman 1985, Rubach 1986



[8] In cases such as the Bulgarian definite article *-at* (masc.) as in *stol* 'chair' – *stol-at* 'the chair', where the *a* is a yer historically and synchronically absent when followed by a gender marker (fem. *-ta*, neut. *-to*), the final empty nucleus of *stol* is also empty in the lexicon. The *a* of the suffix *-at* is floating and attaches to the final empty nucleus, where it is pronounced.

b. Scheer 2004, 2005



In Nsg *són* in (2a), Lower finds that the floating root-internal vowel (a yer) is followed by another floating segment (another yer). Therefore, the leftmost yer is vocalized. This is done by epenthesis of an x-slot (a mora in Hyman's (1985) version) and association of the floating vowel. Epenthesis and association do not occur in Gsg *sn-á* since the root-internal floating vowel is not followed by another floating vowel: the following vowel is the genitive marker *-a*, which does not float. Hence, the floating root-internal vowel remains unassociated and is not pronounced. By contrast, in Gsg *dóm-a* the root-internal vowel is pronounced; it was already associated in the lexicon and therefore does not need epenthesis in order to appear on the surface.

In (2b), the lateral relation that Lower describes is made explicit: government embodies a head-final relationship between a filled and an empty nucleus. It can only be established when the head nucleus is contentful. This is the case in Gsg *sn-á*, where the case marker *-a* fills the final nucleus: the lateral relation goes into effect, and government acts as an association inhibitor that prevents the floating melody of the target nucleus from associating. By contrast, in Nsg *són* the final nucleus is empty and hence no government can be established. Therefore, the root-internal nucleus escapes government and its floating melody can associate. The general rule, then, is that floating melodies can only attach to ungoverned nuclei. Finally, the government-based calculus regarding association is not concerned with stable vowels as in Gsg *dóm-a*, since there are no floating vowels: the case marker *-a* is attached and nothing more happens.

Let us now come back to the issue discussed in section 4.2, namely, that the classical analysis must assume that word-final consonants are followed by yers that never appear on the surface. That is, some yers (those that are word-internal) are sites that host a vowel-zero alternation, while others (those that are word-final) never appear on the surface.

On the analysis in (2b), the two types of yers that are assumed on the classical approach are distinct phonological objects: vowels that alternate with zero are nuclei with a floating piece of melody, while word-final consonants are followed by an empty nucleus (without a floating piece of melody). The Lower computation is not affected by this difference since government is a general device of phonological theory that is active in all languages and therefore describes a relationship between two nuclei—not between two yers.[9]

The issue regarding word-final yers that never appear on the surface thus disappears when it is understood that they are not yers but final empty nuclei. These nuclei are not anything specific

[9] The invariable pattern of vowel-zero alternations across languages and the government-based analysis thereof are described in Scheer 2004:§§15, 411, 426, 2005, 2006, 2011 and Scheer and Ziková 2010.

to yers or Slavic; rather, they are universal in Government Phonology: word-final consonants are onsets of empty nuclei in all languages (see arguments in, for example, Kaye 1990 and Gussmann and Harris 2002). Note that the existence of empty nuclei following word-final consonants is also assumed outside of Government Phonology—for example, by Kiparsky (1991), Burzio (1994), Dell (1995), and Van Oostendorp (2005).

Finally, there is evidence that internal and final yers also show different behavior (Worth 1968, Bethin 1978, Scheer 2012a,b). Russian exhibits variation regarding the vocalization of word-final clusters in Gpl (which, on the classical approach, are followed by a final yer); specifically, some roots do not vocalize (*igr-á – ígr* 'game Nsg, Gpl'), while others do (*kúkl-a – kúkol* 'doll Nsg, Gpl'), and still others offer both options (*igl-á – ígl* / *igól* 'needle Nsg, Gpl'). When followed by an internal yer belonging to a yer-initial suffix such as the diminutive /-ъk/, though, this variation is neutralized: all words obligatorily vocalize (*igór-k-a*, *kúkol-k-a*, *igól-k-a*). This pattern suggests that there is an item following word-final consonants, but that this item is different from word-internal yers.

5 What Is Lexical and What Is Computational

5.1 Patterns Misattributed to Computation

The computational system of a language and the lexicon that contains the items that it manipulates are distinct ontologically, logically, and in the mind as well as in the brain. This distinction is fundamental in linguistics and constitutes a central piece of the generative approach to language: it lies at the heart of the inverted T model where (in production) the morphosyntactic concatenation (today Merge) of pieces (syntactic features) is retrieved from storage and precedes phonological (PF) as well as semantic (LF) interpretation (Chomsky 1965:15–17). Lexical items are stored in long-term memory, while computation is carried out in working (or short-term) memory.

The distinction at hand is also foundational in cognitive science and computational theory (Pylyshyn 1989:55, Newell 1980:156–164), where computational instructions and storage of to-be-computed items are distinct: a process transforms a preexisting object following independently stored instructions.

As mentioned earlier, Becker and Gouskova (2016) propose a hardwired implementation of this distinction: separate grammars control the well-formedness of lexical items (gatekeeper grammar) and input-output computation (grammar proper). On a number of occasions, though, they (see also Gouskova 2012) misattribute to one area what in fact belongs to the other. In the context of Slavic yers, the lexical/computational distinction means that the distribution of yers in lexical items and the workings of the computational mechanism that decides whether lexically present yers appear on the surface are two separate things.

Both Lower and the cluster-based alternative implemented by Yearley (1995) and Gouskova (2012) describe the computational mechanism at hand. This mechanism is not responsible for and has nothing to say about the distribution of yers in the lexical shape of morphemes. As noted earlier, it is undisputed that this distribution cannot be predicted and hence is a lexical accident. But of course the lexical distribution of yers must obey well-formedness requirements applying

to lexical items. For instance, in case there is a synchronically active constraint against unattested *-pk#*, *-tk#* in Russian, there are a number of ways to make sure they cannot occur on the surface. There could be a computational repair deleting either consonant or epenthesizing a vowel in the midst of or after the cluster (grammar proper). Or the sequences could be repaired upon lexicalization (gatekeeper grammar) by the same operations. In case the cluster is broken up by a vowel, this vowel could be stable or a yer. Russian appears to choose the latter option, that is, turning /-pk#/, /-tk#/ into /-p-yer-k#/, /-t-yer-k#/. This has nothing to do with phonological computation (grammar proper).

Building on the case of unattested *-pk#*, *-tk#*, Gouskova (2012:83) says that the fact that "yer alternations . . . are governed by discernible syllable structure constraints" is a well-established generalization. This is a piece of her argument in favor of the cluster-based approach to yer alternations whereby yers are vocalized to avoid certain clusters (rather than because of a following yer, as Lower has it). Her statement is incorrect and wrongly suggests that unattested *-pk#*, *-tk#* clusters play a role in yer vocalization; they do not. The lexical shape of morphemes, not the alternations, is governed by constraints on clusters. In Becker and Gouskova's (2016) approach, these constraints will be active in the gatekeeper grammar (of both the yer and the non-yer lexicon), not in the grammar proper, where input-output mapping is done. Hence, unattested clusters that are broken up by yers cannot be used as an argument for either Lower or the cluster-based approach: they are entirely independent of and irrelevant to the computational mechanism that decides about yer vocalization.

In another passage that confounds the work of the two grammars, Becker and Gouskova (2016:392) suggest that a generalization regarding the lexicon concerns phonological computation (deletion):

> The lexically specific, or unpredictable, nature of yer deletion in individual nouns has long been recognized (Lightner 1965, . . .). Much of this literature debates mechanisms of encoding this lexically specific behavior, or the interactions between yer deletion and other rules of Russian phonology. Nevertheless, more recent work recognizes that there are generalizations about yer deletion that make some predictions possible.

Again, this is meant to be an argument in favor of the cluster-based approach, and against Lower. Becker and Gouskova suggest that the classical literature was not aware of some critical facts regarding the influence of clusters on yer deletion and therefore did not develop the right computational mechanism. This is not a valid argument, though, since the "generalizations about yer deletion" that Becker and Gouskova call on are in fact again generalizations about the lexicon. They note that when yer deletion creates a CCC sequence in Russian, the middle consonant is almost always an obstruent (as in *kost'ór – kostr-á* 'fire Nsg, Gsg') (p. 392); cases with a middle sonorant such as *ágnec – ágnc-a* 'lamb Nsg, Gsg' exist but are very rare. They then show experimentally (through judgment of nonce words) that this lexical proportion is part of the knowledge of speakers since they avoid TRT sequences created by yer deletion.

Becker and Gouskova's (2016) argument is well-taken, but it does not tell us anything about the computational system or yer deletion. Once a TRyerT-V sequence contains a yer, computation will delete it just as in any other sequence, regardless of whether the resulting cluster TRT is

rare, dispreferred, or otherwise costly: it is not ill-formed. Maybe a feedback process involving what speakers know about lexical statistics—that is, lexicon optimization (e.g., Prince and Smolensky 2004 [1993]:225–231, Inkelas 1995)—will at some point eliminate all /TRyerT/ sequences, but this will still be a process concerning the lexicon (hence, encoded in the gatekeeper grammar). Note that the completion of this process does not necessarily mean that the computational system is also affected and becomes unable to produce TRT through yer deletion if such a lexical item were up for computation.

A similar case involves what Becker and Gouskova (2016:393–394) call complex coda blocking (discussed above in section 1). In this case, the authors correctly ascribe the observed phenomenon to the lexicon (a source-oriented generalization in their terms), but they mistakenly suggest that the sequences at hand are ruled out by the gatekeeper grammar ("an offending structure in the base"; p. 394): their own experimental evidence shows that these sequences are dispreferred, but not impossible or ungrammatical.

Becker and Gouskova (2016:393–394) observe that there is not a single word in Russian where a vowel alternates with zero when it is followed by two consonants: in a VCC- sequence, V will never alternate with zero.[10] Their experimental data then support this blocking effect of following clusters: speakers accept deleted vowels significantly more often when they precede root-final singleton consonants (nonce *pišoch – pišch-a*) than when they are followed by a root-

[10] In fact, this generalization is not completely correct: there are two roots whose vowel is followed by an s+C cluster and is present in nominal forms, but absent in verbal and adjectival ones. (These roots are discussed in Gouskova 2012:90–91, but not mentioned in Becker and Gouskova 2016.) That is, *lést'* 'flattery Nsg' alternates with *l'st-ít'*, *l'st-í!*, *l'st-iv-yj* 'to flatter inf., imperative 2sg, fawning adj.' and *mést'* 'revenge Nsg' appears as *mst-ít'* 'to revenge'. Within nominal inflection the root vowel is stable, though: *lést-i* 'flattery Gsg', *mést-i* 'revenge Gsg' (note that the unvocalized imperative *l'st-i!* and the vocalized Gsg *lést-i* ought to represent the same underlying form, /lьst-i/). The fact that these vowels are stable within inflectional paradigms feeds the idea that verbal and nominal forms are not synchronically derived from a common underlier; that is, they do not have independently stored roots (or allomorphs).

A reviewer points out that etymological yers that occurred before a cluster in CS/Old Russian have systematically become stable vowels in Modern Russian: CS **krьst-ъ* > *krést – krest-á* 'cross Nsg, Gsg', CS **dъžd-ь* > *dóžd' – dožd'-á* 'rain Nsg, Gsg', CS **žьzl-ъ* > *žézl - žézl-á* 'mace Nsg, Gsg', CS **čьrv-ь* > *čérv' – čérv'-á* 'worm Nsg, Gsg'. This lends support to the view that Modern Russian vowels before a cluster are always stable, hence that the alternations found in noun-verb pairs are based on two distinct lexical entries (verbal/adjectival /lьst/, nominal /lest/, which may be either different roots or different allomorphs). If they are not stable, a reviewer points out, they require both lexical marking of the underlying form (as having a yer or belonging to the yer lexicon) and an input-output computation that blocks yer vocalization in nouns. Hence, the behavior of the two words in question cannot be captured by a generalization on the source alone.

In sum, the Russian evidence is unlikely to be conclusive regarding the question of whether the synchronic yer-vocalizing mechanism can command vowel-zero alternations before clusters at all. Czech provides better evidence: *lest – lst-i* 'ruse Nsg, Gsg', *mzd-a – mezd* 'wage Nsg, Gpl', and *mst-a – mest* 'revenge Nsg, Gpl' (a fourth case, *čest – ct-i* 'honor Nsg, Gsg', is peculiar in that the vowel alternates but when the zero appears the *s* is lost and the *č* surfaces as *c*, which happens nowhere else in the language). Here, the alternating forms all occur within nominal inflection.

In all cases mentioned, the cluster at stake is *st* or *zd*. This is of course not accidental: s+C sequences display peculiar phonological properties in general—among other things, they are known to behave as if they were a single consonant (Goad 2011). As far as I can see, no Slavic language (at any diachronic stage) has vowel-zero alternations before clusters other than s+C. Items such as the aforementioned CS **čьrv-ь* 'worm' have developed stable vowels in all modern languages.

final cluster (nonce *pišochl – pišchl-a*). What this result shows, though, is precisely that yerCC- roots *are* possible in Russian and that the computation of their yer is perfectly regular. Speakers do not refuse to associate the Gsg *pišchl-a* to the Nsg *pišochl*; they merely make this association less often than when a single consonant separates the yer and the following vowel. In other words, they are able (a) to lexicalize yerCC- items and (b) to have them undergo regular computation, which systematically produces the expected result.

We thus face a generalization about the distribution of yers in the lexicon (there are no morphemes of the shape yerCC-) that is mirrored in speakers' behavior—but this behavior precisely shows that the nonoccurrence of yerCC- items is an accidental gap. Speakers like yerCC- less than yerC-, but they do happily lexicalize and compute items of this shape. Therefore, if speakers come across a neologism or acronym like *Asótr* that someone inflects as *Astr-á*, they will happily lexicalize it as /asъtr/ and subject it to regular phonological computation. This form then fills a gap that, although linguists have detected it in the lexicon, is accidental rather than systematic.

5.2 A Nonquestion

Finally, let us address a question raised by Gouskova (2012:86–89, 100–105): why is it that only the last vowel of a morpheme can alternate with zero in Russian? That is, why is V in CVC items, and V_2 but not V_1 in CV_1CV_2C items, able to alternate with zero? In response to this question, Gouskova (2012:86) says that this phenomenon is "still a mystery," and terms it an asymmetry. She develops a mechanism to derive it, arguing that it must be accounted for by the synchronic computational system of the language.

However, the question has a trivial answer that follows from general principles of how yers work. Devising a computational mechanism to address it is superfluous and again confounds static lexical properties with computation.

The only way to identify a vowel as a yer is to see that it alternates with zero. For a vowel to alternate with zero, its right-hand context must be able to be manipulated: we can see that the vowel of the Russian word for *dream* alternates with zero because we are able to place it both in a context where the following consonant is word-final (result: *són*) and in a context where a vowel follows (result: *sn-á*). Were there no vowel-initial suffixes that could be attached, the root would always be vocalized; and were there no way for /sъn/ to occur word-finally or before a consonant, the yer would always be unpronounced. In neither case would there be grounds for assuming that the vowel is a yer in the first place, either for the linguist or for the child in first language acquisition. In the former case the word would be lexicalized with a stable vowel (/son/: it never appears in any other shape), while in the latter case it would be lexicalized as /sn/ (the only form in which it occurs).

The yer in a hypothetical /CyerCVC/ root would exemplify the latter situation. We would never be able to see the yer on the surface since we are unable to remove the V; hence, the root would always surface as CCVC. As a result, both the analyst and the child would settle on the lexical entry /CCVC/.

A bisyllabic root where both vowels are yers, /CyerCyerC/, is more interesting. There are two patterns in Slavic languages regarding the treatment of yer chains: either all yers in a row are vocalized (save the rightmost), or every other yer is, counting from the right edge. The former pattern is the one described by Lower: Polish /p'ьs-ъk-ъ/ is realized as *pies-ek* 'dog dim. Nsg' (we know independently that both vowels alternate with zero: *ps-a* 'dog Gsg', *pies-k-a* 'dog dim. Gsg'). The latter pattern is the one that Antonín Havlík (1889) discovered in Old Czech and that governs many vowel-zero alternations outside of Slavic (e.g., in Moroccan Arabic, German, and French; see Scheer 2004:§468). It is also active in Old Polish, where /p'ьs-ъk-ъ/ appears as *ps-ek* 'dog dim. Nsg'—that is, without the first yer (other forms are the same as in the modern language).[11]

In a Lower language, a /C_1yerC_2yerC_3/ root will appear on the surface either as $C_1VC_2C_3$ (when the suffix is V-initial: /C_1yerC_2yerC_3-V/) or as $C_1VC_2VC_3$ (elsewhere). That is, the leftmost yer will always appear on the surface; there is no occurrence of the root where it is absent. Therefore, both the analyst and the child will conclude that it is a stable vowel (rather than a yer): the item will be lexicalized as /CVCyerC/. The only evidence that could be brought to bear in order to establish that the leftmost vowel is a yer is diachronic in nature. Gouskova (2012:92) mentions two relevant cases; however, these instantiate not bisyllabic roots but CyerC-ъt-ъ (where -ъ*t* derives abstract nouns): *rópot – rópot-a* 'murmur of discontent Nsg, Gsg' < Common Slavic (CS) **rъp-ъt-ъ* and *tópot* – tópot-a 'tramp of feet Nsg, Gsg' < CS **tъp-ъt-ъ*. Whatever the root structure, the leftmost vowel, etymologically a yer, always appears on the surface as expected (since Russian is a Lower language), and there are no grounds for either the analyst or the child to consider it a yer synchronically; the current lexical items are /ropot/ and /topot/.

By contrast, in a Havlík language the leftmost yer of a /C_1yerC_2yerC_3/ root could alternate with zero. By itself, the root would appear as $C_1C_2VC_3$, (like Old Polish /p'ьs-ъk-ъ/ → *ps-ek*). Followed by a V-initial suffix, however, it would appear as $C_1VC_2C_3$-V. Modern Slavic languages follow the Lower pattern; the Havlík pattern governs earlier diachronic stages. Old Czech, a Havlík language, allows us to ascertain that the pattern whose absence in Russian Gouskova calls "still a mystery" (both V_1 and V_2 of a /CV_1CV_2C/ root alternate with zero) does exist: CS **gъmъz-ъ* 'insects' appears as Old Czech *hmez – hemz-a* 'Nsg, Gsg' and CS **čьbьr-ъ* 'bucket' appears as *čber – čebr-a* 'Nsg, Gsg'.

Hence, Gouskova's observation that the first vowel of a bisyllabic morpheme never alternates with zero does not extend to all Slavic languages. In (Modern) Russian, a Lower language, Gouskova's question is a nonquestion: the analyst and the child necessarily conclude that the first vowel of a bisyllabic morpheme is a stable vowel, not a yer. This is a fact about the lexicon and

[11] On the difference between the Lower pattern and the Havlík pattern, see Bethin 1998:209–214 and Scheer and Ziková 2010. On all accounts, the contrast lies in the cyclic (Lower) vs. noncyclic (Havlík) application of the yer-vocalizing mechanism.

has nothing to do with the computational system of the language. Devising a computational mechanism to exclude yers from the position of first vowels in bisyllabic morphemes confuses lexical distribution and computation.

6 Conclusion

As the title indicates, the main purpose of this reply is to clarify where exactly the red line runs between the computational mechanism that vocalizes yers and their lexical distribution in morphemes. On a number of occasions, these are confounded by Gouskova (2012) and Becker and Gouskova (2016), for example, when Gouskova (2012) devises a computational mechanism to account for the lexical distribution of yers (section 5.2). In the same way, when talking about unattested *-pk#*, *-tk#* in Russian, it is incorrect to say that "yer alternations . . . are governed by discernible syllable structure constraints" (Gouskova 2012:83); and, when talking about the generalization that the middle consonant of a yer-deletion-created CCC sequence is almost always an obstruent, it is incorrect to say that "more recent work recognizes that there are generalizations about yer deletion that make some predictions possible" (Becker and Gouskova 2016:392) (section 5.1).

Another aspect of the lexical shape of morphemes is the distinction between systematic and accidental gaps. Given that yers do not occur before morpheme-final clusters in Russian, Becker and Gouskova's (2016) nonce-word-based experiments show that speakers also prefer items that are distinct from yerCC-. They argue that *yerCC- is thus part of speakers' knowledge, and this is certainly correct. However, their evidence shows precisely that the absence of yerCC- items in Russian is an accidental gap: speakers like yerCC- less than yerC-, but they do happily lexicalize and compute items of this shape. Hence, we are talking about a tendency that is active in the lexicalization of items, not about a systematic prohibition.

Finally, Becker and Gouskova (2016) report a preference for yer-deletion-created (C)TR clusters when speakers process nonce words. However, sonority patterns in clusters that result from yer deletion are irrelevant (or in fact dysfunctional) for Russian grammar (lexicon, input-output computation, acquisition) (section 2). The sonority sensitivity in Becker and Gouskova's experimental data may stem from a universal bias in favor of TR clusters (Berent 2013) that is independent of Russian.

Given these results, what have we learned about whether yer vocalization is cluster-based (Gouskova 2012, Becker and Gouskova 2016) or cluster-independent (Lower)? On a number of occasions, Becker and Gouskova (2016) argue that Lower was established at a time when relevant cluster-based generalizations had not yet been discovered, hence misses out on relevant empirical material that invalidates its central idea that clusters are irrelevant for yer vocalization. However, this argument is ill-founded since the phenomena it is based on are not computational (input-output) in nature, as Becker and Gouskova mistakenly hold. That clusters belong to the lexicon confirms their irrelevance for yer vocalization and hence supports Lower (section 5.1). Moreover, cluster-based yer vocalization assumes that vowel-zero alternations are in fact five different processes with different causalities and hence misses the fact that the process follows a single and unified rationale. Lower, on the other hand, offers a unified analysis based on one single cause

(section 4.1). Finally, the major objection raised against Lower (final yers that never appear on the surface) disappears under the government-based implementation of Lower (section 4.3).

Whatever the correct solution for yer vocalization turns out to be, how does the preceding discussion speak to the question whether, using Gouskova's terms, morphemes or segments (yers) are exceptional? Let us first look at the computational/lexical distinction. Once distributional patterns in the lexicon and input-output computation are properly distinguished, one could take either view regarding what yer vocalization refers to (at least in Russian, where morphemes contain at most one yer): morphemes marked as yer-containing or segments marked as yers. Also, following Becker and Gouskova (2016), it seems reasonable that there is a mechanism enforcing (or inducing a preference for) certain shapes of morphemes upon lexicalization (*SPE*'s morpheme structure constraints, Becker and Gouskova's gatekeeper grammar).[12]

There is a logical dependency between cluster-based vocalization and the morpheme-based view, though. As shown in section 3, the cluster-based analysis of yer vocalization cannot distinguish CC# clusters that are broken up from CC# clusters that are not; therefore, it needs to assign those that are broken up to a yer sublexicon and those that are not to a non-yer sublexicon. That is, the cluster-based analysis is only viable with morpheme-based yer vocalization. Segment-based vocalization could not make reference to clusters: there would be no reason to vocalize /-CyerC#/ sequences if nonvocalization produces a legal surface CC# cluster (as in /lasъk/ → *lások* 'weasel Gpl' in the presence of /lask/ → *lásk* 'caress Gpl').

On the other hand, although the morpheme-based identification of yers can work with cluster-based vocalization, it is also compatible with the classical implementation of Lower. This is because in this implementation, two consecutive yers never belong to the same morpheme (at least in Russian, where morphemes contain at most one yer); recall from section 4.2 that yers that occur after word-final consonants are interpreted as case markers (-ъ in /sъn-ъ/ → *són* marks Nsg). Lower could therefore be stated as "Vocalize a yer if and only if the following morpheme contains a yer" (instead of " . . . if and only if the following vowel is a yer"). The question is why this should be done: there is no import with respect to the regular formulation, and it seems odd for a phonological process to make reference to morphology when reference to phonology alone can do the job and avoid setting up different sublexicons.

Things are different, though, regarding the alternative implementation of Lower discussed in section 4.3. Unlike in the classical implementation, the yer of *són* does not appear on the surface because it is followed by another yer that represents Nsg; rather, it appears because it is followed by an empty nucleus (/sъnØ/ → *són*). Hence, on this analysis yer vocalization cannot be described as occurring " . . . if and only if the following morpheme contains a yer": vocalization occurs if and only if the following nucleus is not associated to any melody, and the triggering empty nucleus in /sъnØ/ belongs to the same morpheme as the vocalized yer.

[12] In an Optimality Theory environment, the compatibility of such a device with Richness of the Base begs the question, though.

In sum, then, the situation is as follows. If it can be shown that cluster-based vocalization is incorrect, the necessity for morpheme-based yer identification disappears (the cluster-based option requires morpheme-based yer identification). And indeed, as shown in section 4.1, the cluster-based analysis misses the basic insight into Slavic vowel-zero alternations and fails to provide a unified account for a unified phenomenon. We are thus left with the logical possibility that identification of yers can be morpheme-based when vocalization is driven by Lower in its classical implementation—an option that is, however, costly and unnecessary. On the other hand, if yer vocalization is triggered by empty nuclei, morpheme-based yer identification is incompatible with Lower.

References

Becker, Michael, and Maria Gouskova. 2016. Source-oriented generalizations as grammar inference in Russian vowel deletion. *Linguistic Inquiry* 47:391–425.

Berent, Iris. 2013. *The phonological mind.* Cambridge: Cambridge University Press.

Berent, Iris, Tracy Lennertz, Jongho Jun, Miguel A. Moreno, and Paul Smolensky. 2008. Language universals in human brains. *Proceedings of the National Academy of Sciences of the United States of America* 105:5321–5325.

Bethin, Christina. 1978. Phonological rules in the nominative singular and the genitive plural of the Slavic substantive declension. Doctoral dissertation, University of Illinois at Urbana-Champaign.

Bethin, Christina. 1998. *Slavic prosody: Language change and phonological theory.* Cambridge: Cambridge University Press.

Burzio, Luigi. 1994. *Principles of English stress.* Cambridge: Cambridge University Press.

Chomsky, Noam. 1965. *Aspects of the theory of syntax.* Cambridge, MA: MIT Press.

Chomsky, Noam, and Morris Halle. 1968. *The sound pattern of English.* New York: Harper and Row.

Czaykowska-Higgins, Ewa. 1988. *Investigations into Polish morphology and phonology.* Cambridge, MA: MIT Press.

Dell, François. 1995. Consonant clusters and phonological syllables in French. *Lingua* 95:5–26.

Do, Youngah, Chiyuki Ito, and Michael Kenstowicz. 2014. The base of Korean noun paradigms: Evidence from tone. *Korean Linguistics* 16:111–143.

Dresher, Elan. 2009. *The contrastive hierarchy in phonology.* Cambridge: Cambridge University Press.

Farina, Donna. 1991. Palatalization and jers in modern Russian phonology: An underspecification approach. Doctoral dissertation, University of Illinois at Urbana-Champaign.

Goad, Heather. 2011. The representation of sC clusters. In *The Blackwell companion to phonology*, ed. by Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume, and Keren Rice, 898–923. New York: Wiley-Blackwell.

Gouskova, Maria. 2012. Unexceptional segments. *Natural Language and Linguistic Theory* 30:79–133.

Gouskova, Maria, and Michael Becker. 2013. Nonce words show that Russian yer alternations are governed by the grammar. *Natural Language and Linguistic Theory* 31:735–765.

Gussmann, Edmund. 1980. *Studies in abstract phonology.* Cambridge, MA: MIT Press.

Gussmann, Edmund. 2007. *The phonology of Polish.* Oxford: Oxford University Press.

Gussmann, Edmund, and John Harris. 2002. Word-final onsets. In *UCL working papers in linguistics 14*, ed. by Ad Neeleman and Reiko Vermeulen, 1–42. London: University College London, UCL Division of Psychology and Language Sciences.

Hall, Daniel Currie. 2011. Phonological contrast and its phonetic enhancement: Dispersedness without dispersion. *Phonology* 28:1–54.

Havlík, Antonín. 1889. K otázce jerové v staré češtině. *Listy Filologické* 16:45–51, 106–116, 248–258, 342–353, 436–445.

Hyman, Larry. 1985. *A theory of phonological weight.* Dordrecht: Foris.
Inkelas, Sharon. 1995. The consequences of optimization for underspecification. In *NELS 25*, ed. by Jill N. Beckman, 287–302. Amherst: University of Massachusetts, Graduate Linguistic Student Association. (Also Rutgers Optimality Archive 240, http://roa.rutgers.edu.)
Ito, Chiyuki. 2010. Analogy and lexical restructuring in the development of nominal stem inflection from Middle to Contemporary Korean. *Journal of East Asian Linguistics* 19:357–383.
Kaye, Jonathan. 1990. 'Coda' licensing. *Phonology* 7:301–330.
Kaye, Jonathan, Jean Lowenstamm, and Jean-Roger Vergnaud. 1990. Constituent structure and government in phonology. *Phonology* 7:193–231.
Kenstowicz, Michael, and Jerzy Rubach. 1987. The phonology of syllabic nuclei in Slovak. *Language* 63: 463–497.
Kiparsky, Paul. 1991. Catalexis. Ms., Stanford University, Stanford, CA.
Laskowski, Roman. 1975. *Studia nad morfonologią współczesnego języka polskiego.* Wrocław: Wydawnictwo Polskiej Akademii Nauk.
Lightner, Theodore. 1965. Segmental phonology of Contemporary Standard Russian. Doctoral dissertation, MIT, Cambridge, MA.
Lowenstamm, Jean. 1996. CV as the only syllable type. In *Current trends in phonology: Models and methods, vol. 2*, ed. by Jacques Durand and Bernard Laks, 419–441. Salford, Manchester: ESRI.
Łukaszewicz, Beata. 2006. Extrasyllabicity, transparency and prosodic constituency in the acquisition of Polish. *Lingua* 116:1–30.
Newell, Allen. 1980. Physical symbol systems. *Cognitive Science* 4:135–183.
Oostendorp, Marc van. 2005. The first person singular in Dutch dialects. In *NELS 35*, ed. by Leah Bateman and Cherlon Ussery, 1–12. Amherst: University of Massachusetts, Graduate Linguistic Student Association.
Pater, Joe. 2009. Morpheme-specific phonology: Constraint indexation and inconsistency resolution. In *Phonological argumentation: Essays on evidence and motivation*, ed. by Steve Parker, 123–154. London: Equinox.
Pesetsky, David. 1979. Russian morphology and lexical theory. Ms., MIT, Cambridge, MA. http://web.mit.edu/linguistics/www/pesetsky/russmorph.pdf.
Piotrowski, Marek. 1992. Polish yers in non-linear phonology. In *Phonologica 1988*, ed. by Uli Dressler, Hans Luschützky, Oskar Pfeiffer, and John Rennison, 215–227. Cambridge: Cambridge University Press.
Prince, Alan, and Paul Smolensky. 2004 [1993]. *Optimality Theory: Constraint interaction in generative grammar.* Oxford: Blackwell.
Pylyshyn, Zenon. 1989. Computing in cognitive science. In *Foundations of cognitive science*, ed. by William Posner, 51–91. Cambridge, MA: MIT Press.
Rubach, Jerzy. 1984. *Cyclic and lexical phonology: The structure of Polish.* Dordrecht: Foris.
Rubach, Jerzy. 1986. Abstract vowels in three dimensional phonology: The yers. *The Linguistic Review* 5:247–280.
Rubach, Jerzy. 1993. *The lexical phonology of Slovak.* Oxford: Clarendon Press.
Rubach, Jerzy. 2013. Exceptional segments in Polish. *Natural Language and Linguistic Theory* 31: 1139–1162.
Scheer, Tobias. 2004. *A lateral theory of phonology.* Vol. 1, *What is CVCV, and why should it be?* Berlin: Mouton de Gruyter.
Scheer, Tobias. 2005. Slavic vowel-zero alternations and Government Phonology: Two approaches, one solution. In *Formal Approaches to Slavic Linguistics 13: The South Carolina Meeting*, ed. by Steven Franks, Frank Gladney, and Mila Tasseva-Kurktchieva, 300–311. Ann Arbor: Michigan Slavic Publications.

Scheer, Tobias. 2006. How yers made Lightner, Gussmann, Rubach, Spencer and others invent CVCV. In *Studies in constraint-based phonology*, ed. by Piotr Bański, Beata Łukaszewicz, and Monika Opaliń-ska, 133–207. Warsaw: Wydawnictwo Uniwersytetu Warszawskiego.

Scheer, Tobias. 2011. Slavic yers. In *The Blackwell companion to phonology*, ed. by Marc van Oostendorp, Colin Ewen, Elizabeth Hume, and Keren Rice, 2936–2962. New York: Wiley-Blackwell.

Scheer, Tobias. 2012a. Variation is in the lexicon: Yer-based and epenthetic vowel-zero alternations in Polish. In *Sound, structure and sense: Studies in memory of Edmund Gussmann*, ed. by Eugeniusz Cyran, Bogdan Szymanek, and Henryk Kardela, 631–671. Lublin: Wydawnictwo KUL.

Scheer, Tobias. 2012b. Yers and epenthetic vowels in Polish. In *Slavic languages in formal grammar*, ed. by Mojmir Dočekal and Markéta Ziková, 179–204. Frankfurt am Main: Lang.

Scheer, Tobias, and Markéta Ziková. 2010. The Havlík pattern and directional Lower. In *Formal Approaches to Slavic Linguistics 18: The Second Cornell Meeting 2009*, ed. by Draga Zec and Wayles Browne, 471–486. Ann Arbor: Michigan Slavic Publications.

Szpyra, Jolanta. 1992. Ghost segments in nonlinear phonology: Polish yers. *Language* 68:277–312.

Szpyra, Jolanta. 1995. *Three tiers in Polish and English phonology.* Lublin: Wydawnictwo Universytetu Marii Curie-Skłodowskiej.

Townsend, Charles E. 1975. *Russian word-formation.* Columbus, OH: Slavica.

Worth, Dean. 1968. Vowel-zero alternations in Russian derivation. *International Journal of Slavic Linguistics and Poetics* 11:110–123.

Yearley, Jennifer. 1995. Jer vowels in Russian. In *Papers in Optimality Theory*, ed. by Jill N. Beckman, Laura Walsh Dickey, and Suzanne Urbanczyk, 533–571. Amherst: University of Massachusetts, Graduate Linguistic Student Association.

Université Côte d'Azur, CNRS, Bases Corpus Langage (BCL)
Pôle universitaire Saint Jean d'Angély 3
24 avenue des Diables Bleus
F-06357 Nice Cedex 4
France

tobias.scheer@unice.fr