



What Are Constructions, and What Else Is Out There? An Associationist Perspective

Vsevolod Kapatsinski*

Department of Linguistics, University of Oregon, Eugene, OR, United States

OPEN ACCESS

Edited by:

Matthew Carlson,
Pennsylvania State University (PSU),
United States

Reviewed by:

Laura A. Janda,
Arctic University of Norway, Norway
Sara Finley,
Pacific Lutheran University,
United States

*Correspondence:

Vsevolod Kapatsinski
vkapatsi@uoregon.edu

Specialty section:

This article was submitted to
Language Sciences,
a section of the journal
Frontiers in Communication

Received: 22 June 2020

Accepted: 10 December 2020

Published: 07 January 2021

Citation:

Kapatsinski V (2021) What Are
Constructions, and What Else Is Out
There? An Associationist Perspective.
Front. Commun. 5:575242.
doi: 10.3389/fcomm.2020.575242

Constructionist approaches to language propose that the language system is a network of constructions, defined as bidirectional mappings between a complex form and a meaning. This paper critically evaluates the evidence for and against two possible construals of this proposal as a psycholinguistic theory: that direct, bidirectional form-meaning associations play a central role in language comprehension and production, and the stronger claim that they are the only type of association at play. Bidirectional form-meaning associations are argued to be plausible, despite some apparent evidence against bidirectionality. However, form-meaning associations are insufficient to account for some morphological patterns. In particular, there is convincing evidence for productive paradigmatic mappings that are phonologically arbitrary, which cannot be captured by form-meaning mappings alone, without associations between paradigmatically related forms or constructions. Paradigmatic associations are argued to be unidirectional. In addition, subtraction and backformation at first glance require augmenting the associative networks with conditioned operations (rules). However, it is argued that allowing for negative form-meaning associations accommodates subtraction and backformation within the constructionist approach without introducing any additional mechanisms. The interplay of positive and negative form-meaning associations and paradigmatic mappings is exemplified using a previously undescribed morphological construction in Russian, the *bez*-Adjective construction.

Keywords: construction grammar, usage-based linguistics, associative learning, morphology, morphological paradigms, second-order schemas, Russian

INTRODUCTION

In constructionist approaches to language, the grammar and the lexicon are unified into the constructicon, a single network of constructions, defined as form-meaning pairings (e.g., Bybee, 1985; Langacker, 1987; Goldberg, 1995, 2006; Kapatsinski, 2013, 2014; Diessel, 2015). All languages have constructica, and the ability to acquire a large constructicon is the crucial pre-requisite to acquiring a human language (Deacon, 1997). Constructions are agreed to be probabilistic, multiply determined and learned by generalization over experienced utterances.

However, two issues remain unresolved. First, an important issue involves directionality: are constructions really bidirectional form-meaning mappings, Saussurean signs, or are there separate sets of form→meaning and meaning→form mappings? While the former has been the default assumption in the literature, there are also strong arguments for assuming otherwise (e.g., Ramscar et al., 2010). Second, there is disagreement on whether form-meaning mappings are sufficient

to explain utterance comprehension and production, or if the mental grammar also contains other types of generalizations. One such generalization type is represented by second-order, paradigmatic generalizations mapping one construction onto another. Paradigmatic or “second-order” generalizations can be of two kinds. Within the constructionist framework, the grammar is a network of mappings, and second-order generalizations are paradigmatic mappings between constructions (Ford et al., 1997; Cappelle, 2006; Nessel, 2008; Booij, 2010; Kapatsinski, 2013, 2017b, 2018; Booij and Audring, 2017a,b; Audring, 2019). In the generative framework, second order generalizations are rules (context-specific operations) that transform a base construction into another one, in a certain context (Albright and Hayes, 2003; Kapatsinski, 2010a). As pointed out by Pinker and Prince (1988), operations are not consistent with an associationist approach to the mind and require an additional mechanism.

Paradigmatic generalizations of both kinds have been explicitly questioned in the constructionist literature (e.g., Bybee, 2001; Goldberg, 2002) because there is less need for them than in a framework that does not posit constructions. Indeed, I will argue below that we do not have evidence that they are necessary above the word level (i.e., in syntax). However, morphology provides crucial evidence for the existence of paradigmatic mappings and/or rules (Nessel, 2008; Booij, 2010; Becker and Gouskova, 2016; Booij and Audring, 2017a,b). I argue that rules may not be necessary if paradigmatic mappings are allowed and associations can be inhibitory.

The workings of an associative network are illustrated using a previously undescribed construction in Russian, the *bez* Adjective construction. A full description of the construction requires us to make use of all types of associations: *schematic* associations between meanings and forms, as well as syntagmatic¹ and paradigmatic associations between forms, and requires both excitatory and inhibitory schematic associations. It also illustrates two fundamental but less controversial properties of an associative network: that multiple bases are used to produce a novel wordform via a multitude of parallel routes.

ARE BIDIRECTIONAL FORM-MEANING ASSOCIATIONS (CONSTRUCTIONS) POSSIBLE?

Constructions are typically defined as pairings between form and meaning, a definition that brings with it at least an implicit assumption of bidirectionality. In an associative network, bidirectional mappings² mean that activation of a form changes activation of a meaning as much as activation of the meaning changes activation of the form. One worries about at least two

ways in which this assumption may not hold water. First, the connection from a form to a meaning might have a different strength than the connection from the meaning to the form. Second, there may not even be a single form level used for production and comprehension: if we take the constructionist assumption of there being no levels of abstraction between meaning and form to its logical extreme, then the forms in comprehension of a spoken language would be auditory or audiovisual in nature while the forms in production would be articulatory. In contrast, bidirectionality requires a form level that mediates the mapping from audition to semantics in comprehension and from semantics to articulation in production (Kapatsinski, 2018, pp.59-62).

Production-Comprehension Dissociations Are Predicted by Bidirectional Associations

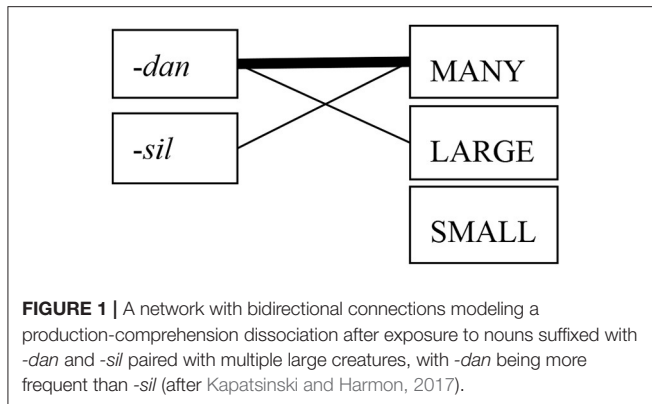
The existence of bidirectional form-meaning associations can be questioned on the ground that there exist production/comprehension dissociations. In particular, when multiple forms are competing to express a meaning, the form most likely to be chosen to express the meaning in production may not be the form that would best transmit the meaning to the listener, i.e., the form that is the best cue to that meaning in comprehension (Kapatsinski, 2012; Harmon and Kapatsinski, 2017; Koranda et al., 2018). Production-comprehension dissociations of this type can even be observed within the same individual. However, the existence of such dissociations does not necessarily imply that the connections between form and meaning are unidirectional.

The basic reason that production-comprehension dissociations are not probative regarding directionality of form-meaning connections is that there is always a reason for choosing a form i.e., not the best cue to the meaning to be expressed. Often, this reason can be incorporated into the network as an additional cue to the form, which contributes to form choice independently of the bidirectional form-meaning connection. To take an extreme example, a bilingual chooses among cognate constructions in part based on how strongly they are activated by the meaning to be expressed, but also based on the language that the listener is likely to understand.

An additional reason for dissociations is the role of form accessibility in production choice. Frequent, more accessible words can be chosen for production over less frequent alternatives even when those infrequent alternatives would be better cues to the meaning for the listener. As discussed by Harmon and Kapatsinski (2017), Harmon (2019), and Smolek (2019), this mechanism can explain regularization and paradigm leveling in language change. Harmon and Kapatsinski (2017) show experimentally that increasing the frequency of a construction in a learner's experience makes the construction's form more likely to be used to express *related* meanings, and yet makes the same speaker more confident that it does not map onto these meanings in comprehension. For example, learners who experience the constructions *N-dan* and *N-sil* paired with multiple large creatures a few times are equally likely to pick multiple small creatures and multiple large creatures in response

¹Syntagmatic generalizations could be used to predict upcoming forms and retrodict the forms one has missed (Lieberman, 1963; Osgood, 1963). An alternative is posed by interactive activation flow between parts and wholes (McClelland and Elman, 1986). Because of space restrictions, I will simply assume syntagmatic associations here.

²*Mappings* may be implemented by multiple *associations* in the mind, or by multiple *connections* in a neural network.



to either form. However, as they keep encountering *N-dan* with multiple large creatures they stop selecting multiple small creatures in response to examples of *N-dan*. They become more confident that “large” is a necessary part of the meaning of *-dan* (see also Xu and Tenenbaum, 2007). However, they also become more likely to choose *-dan* to name multiple small creatures. Koranda et al. (2018) also demonstrate this effect in a continuous semantic domain of angles: learners use frequent terms to refer to a broader range of angles although they are more confident about what angle a frequent term actually refers to.

Accessibility-driven dissociations are easily modeled with bidirectional form-meaning connections, as shown in **Figure 1**. We assume that connections between forms and meanings strengthen when the form is paired with the meaning, that *MANY* (i.e., PLURAL) is a more salient meaning for an English speaker than *LARGE*, and that connections between salient cues and outcomes develop faster than connections between less salient ones. When the listener is presented with *-dan*, s/he knows that it means *LARGE* rather than *SMALL*: *-dan* activates *LARGE* more than *SMALL* because it has grown an association with the (initially low-salience *LARGE* feature). When presented with *-sil*, the listener activates *MANY* but has no way to pick between *LARGE* and *SMALL*. Because of the initially low salience of *LARGE* as a feature to an English speaker, the small number of exposures to *-sil* paired with *LARGE* has not allowed that connection to develop. As a speaker, the same participant will choose *-dan* over *-sil* when presented with *MANY SMALL* because there is no connection from *SMALL* to either form, while the *MANY-dan* connection is stronger than the *MANY-sil* connection. Thus, production-perception dissociations of this kind are actually predicted by bidirectional form-meaning associations.

Bidirectional Associations Can Be Learned With Unidirectional Mechanisms

A different kind of argument against bidirectionality was presented by Ramsar et al. (2010). Ramsar et al. presented participants with training trials featuring form-meaning pairings, with meanings being visual depictions of novel 3D objects. The crucial manipulation was whether the form preceded the meaning or the meaning preceded the form. The meanings were

clustered into six categories, with two categories paired with each of the words *wug*, *tob* or *dep*. There were salient non-discriminative visual features that distinguished subcategories paired with the same form, but were shared between forms, and therefore would not allow the learner to predict the form from the object. There were also non-salient discriminative visual features that defined only one of the subcategories paired with a form but were not shared between forms, and would therefore allow the learner to predict forms. Learners in the meaning-before-form condition picked up on these low-salience discriminative features, while those in the form-before-meaning condition did not.

Ramsar et al. (2010) argue that learners acquire meaning-to-form connections following the classic Rescorla-Wagner learning rule, which uses sets of cues to predict upcoming outcomes. The Rescorla-Wagner rule updates cue-to-outcome connection weights in proportion to the unexpectedness of the outcome's occurrence, or absence (Rescorla and Wagner, 1972), as well as cue and outcome salience. The rule is unidirectional in two senses: it does not learn outcome-to-cue associations, and the roles of cues and outcomes during learning are different. The rule assumes that the learner predicts whether the outcome will occur based on the present cues. Thus, the learner has expectations about which outcomes would occur in various contexts and can learn when those outcomes are *unexpectedly* absent. The learner does not form expectations about the contexts in which cues will occur, and therefore learns nothing from absent cues, and cue salience is due solely to its inherent perceptual properties. Because of these asymmetries in the rule, Ramsar et al. conclude that “the relationship between symbols and the things they represent is not bidirectional” (p.912).

The results of Ramsar et al. provide an important illustration of the importance of prediction in language learning, and document the existence of cue competition between semantic cues. The results are indeed consistent with the predictions of the Rescorla-Wagner rule, and support a different role for cues and outcomes during learning. Nonetheless, learning may result in associations being formed in both directions, allowing the learner to predict an outcome based on a cue and to infer that a missed cue must have occurred based on an outcome. Bidirectionality would imply that the associations in both directions should be equal in strength. Ramsar et al.'s results are consistent with this prediction. In their experiment, participants are tested both on choosing forms given a meaning and on choosing a meaning given a form (p.930). Participants who experienced the meaning-before-form condition appear to have performed better on both tasks (accuracy within task is not reported, and no task differences are reported). It therefore appears that form-to-meaning connections benefit from the meaning-before-form order as much as meaning-to-form connections do. This result is *prima facie* inconsistent with participants learning meaning-to-form connections in one condition and form-to-meaning connections in the other; strong correlations between the strengths of $A \rightarrow B$ and $B \rightarrow A$ are a classic argument for bidirectionality (Kahana, 2002). Rather, the results are consistent with the alternative hypothesis that participants in the meaning-before-form condition are learning bidirectional or reciprocal

associations between discriminative semantic features and the forms they predict. That is, cues that are predictive of outcomes acquire salience and are associated with those outcomes more strongly than cues that are not predictive, but the associations themselves appear to form in both directions.

Second, Ramscar et al. (2010, p.918) assume that forms do not have identifiable subparts that can compete to predict meanings (though cf. Blevins et al., 2016):

“verbal labels are relatively discrete and possess little cue structure [...] consider the label *pan*. A native English speaker can parse it into a sequence of phonemes [p^han] but will be largely unable to discriminate further cues within these sounds [...]. Because phonemes are perceived sequentially rather than simultaneously [...] phonemes cannot compete directly as cues. Moreover, the other discriminable cues present in speech—such as emphasis, volume, and pitch contour—do not covary systematically with phonemes.”

This paragraph above denies both the existence of subphonemic cues, including phonetic cues and phonological features, and the pervasiveness of coarticulation. Phonemes do of course covary with loudness or pitch; for example, pitch at the beginning of a vowel is a secondary cue to the voicing of the preceding stop, distinguishing [p] from [b], and listeners are exquisitely sensitive to such patterns of covariation (e.g., Idemaru and Holt, 2011). Likewise, it is not the case that the cues to phonemes are strictly sequential. For example, the place cues to the [p] are in the formants also identifying the height and frontness of the following vowel. Given these considerations, we should expect that phonetic cues compete with each other for predicting meanings, and indeed this effect has recently been documented by Nixon (2020).

Third, the way that children experience forms and objects is often very different from the experimental setup in Ramscar et al. (2010), where the object was presented very briefly (for 150 ms) either a second after or a second before the spoken word. Head camera data indicate that most efficacious word learning episodes involve parents naming objects that the child is already looking at, and the child continuing to look at the object during the label and for some time afterwards (Pereira et al., 2014). These experiences allow the child to both predict the form from the meaning, and to predict meaning from the form, potentially forming a connection or connections in both directions.

Finally, in conditioning experiments that have provided the motivation for the Rescorla-Wagner model, cues are predictive but devoid of inherent value whereas the outcomes are biologically relevant events like the dispensation of food or electric shock. Because of this inherent asymmetry, it is plausible that cues are used to predict outcomes and not vice versa (although research in animal learning has argued for models incorporating reciprocal cue-outcome connections; Matzel et al., 1988; Honey et al., 2020). Because wordforms and other speech sounds are not themselves biologically relevant, while the events they predict often are—especially in the early experience of an infant—it appears implausible to restrict learners to predicting forms from meanings.

In conclusion, there is no current empirical evidence against bidirectionality of form-meaning mappings. Dissociations between production and comprehension can be observed, but are predicted from simple models with bidirectional associations. There is evidence that such associations can be learned by predicting forms from meanings, but the learned associations can then be used to select meanings given forms as well as to select forms given meanings.

What Are “Form” Representations?

While bidirectionality is consistent with the behavioral evidence, it does raise questions about how it could be implemented in the brain. On the one hand, many brain areas are connected bidirectionally: there is just as much top-down activation flow (from meanings to forms) as bottom-up activation flow (see O’Reilly and Munakata, 2000, for an excellent review), a fact that has provided a motivation for interactive activation models of language processing (Dell, 1986; McClelland and Elman, 1986; O’Reilly and Munakata, 2000) and Grossberg’s Adaptive Resonance Theory (Grossberg, 1987, 2013). While the top-down and bottom-up connections largely involve separate neurons, it is not impossible to imagine bidirectionality at the level of forms and meanings of constructions, which correspond to activation patterns distributed over large populations of neurons (see Allen et al., 2012, for an attempt to identify such patterns in fMRI). For example, in Grossberg’s theory, constructions would be resonant brain states in which the form level and the meaning level feed activation to each other, helping maintain a construction in an activated state for the significant period of time likely necessary for constructions to guide utterance planning. Based on Pereira et al. (2014) head camera data, efficacious naming episodes tend to provide children with the opportunity to establish such a resonance as an object persists in the child’s view before, during and after the referring form is heard.

The bidirectional top-down and bottom-up activation flows connect semantic representations to perceptual and (pre)motor processing areas of the brain. However, to say that the same form-meaning mappings are active in comprehension and production requires the two processing directions to share a form level. The need for a form level appears to preclude a radical exemplar account of language (e.g., Ambridge, 2019) in which there is no significant abstraction, and therefore constructions map perceptual representations onto meanings in comprehension and meanings onto motor representations in production. The question of whether there is a level of form representations shared between perception and production has been a long-standing area of debate in phonetics. A promising direction for unifying the two is represented by Bayesian analysis-by-synthesis models, in which the listener evaluates hypotheses about possible production representations that could have generated the perceived auditory signal (e.g., Bever and Poeppel, 2010). If these accounts are on the right track, then production representations could serve as the form level mediating between audition and semantics in speech perception, and bidirectional form-meaning connections could connect these production representations to meanings. Literal bidirectionality could also

be maintained by models in which the production targets the speaker aims to achieve are perceptual in nature.

But what if there is no form level? What if listeners map perceptual representations directly onto meanings, while speakers map meanings directly onto articulatory representations? In this case, the mappings would not be constrained to be bidirectional by the architecture of the language system. However, I would argue, learning would nonetheless modify the weights of the unidirectional mappings to bring them in close alignment, allowing for bidirectionality in activation flow. Models that posit separate form levels for production and perception also posit mechanisms for bringing those form levels into alignment during early development (Guenther and Perkell, 2004; Davis and Redford, 2019). Through these mechanisms, which likely include both reinforcement learning and imitation, production and comprehension representations appear to be linked so closely that activating one appears to of necessity activate the other. Although there is debate regarding whether the motor cortex plays a *mediating* role in speech perception, there is consensus that it is *activated* by speech sounds. Likewise, there is recent evidence that silent speech produces activation in the auditory cortex (Okada et al., 2018). If the two form levels necessarily activate and resonate with each other in both production and comprehension, the linked representations function as a single form level i.e., both activated by and activates semantics. That is, if articulatory and perceptual representations necessarily activate each other, it is possible for a meaning to always increase activation of an articulatory representation to the same extent that the corresponding perceptual representation increases activation of the meaning, allowing for bidirectionality.

WHEN ARE SECOND-ORDER GENERALIZATIONS NEEDED?

Usage-based constructionist approaches to grammar are skeptical of transformations, and question the need to derive constructions from either other constructions or underlying forms (e.g., Bybee, 1985, 2001; Langacker, 1987; Goldberg, 2002; Diessel, 2015). However, it has been argued that there are second-order schemas relating “allostructions” (Cappelle, 2006) or, more generally, constructions that share parts (Ford et al., 1997; Nessel, 2008; Booij, 2010; Kapatsinski, 2013, 2017b, 2018; Jackendoff and Audring, 2016; Booij and Audring, 2017b; Audring, 2019). In syntax, for example, Cappelle (2006) has argued that there is a need to relate the English verb-particle-NP construction to the verb-NP-particle construction, as in *When did you give it up* vs. *When did you give up drinking*. In morphology, there is apparent need to relate words that share a stem. For example, *ambition* and *ambitious*, *caution* and *cautious* can be related together by a schema linking together the [...ous]_A and [...ion]_N constructions, which would encode the fact that an adjective ending in *-ous* usually corresponds to a noun ending in *-ion* and not some other nominal suffix (Audring, 2019). Another well-known example is the [...ist]~[...ism] schema as in *pacifist*~*pacifism*, which allows one to explain how one would

derive an *-ist* adjective from a new *-ism* noun or vice versa (Booij, 2010). As Booij pointed out, these kinds of direct mappings capture the fact that the semantic relationship between the *-ist* and *-ism* forms is regular whereas this cannot be said of each form's relation to its stem. An X-ist can have many semantic relationships to X, but necessarily believes in X-ism.

Second-Order Schemas Are Rare but Necessary for Morphology

Jackendoff and Audring (2016) have argued that second-order schemas are ubiquitous, and that any two constructions that share some aspect of form or meaning are linked by a second-order schema. Furthermore, second order schemas can be posited even if they are not productive. However, from a usage-based perspective, a generalization plays a role in the grammar if it is *used* to understand or produce language, thus unproductive schemas are rather suspect. Dabrowska and Szczerbiński (2006) and Engelmann et al. (2019), among others, show that many speakers of highly inflected languages may not use many of the second-order schemas of their language productively, with productivity of a schema being a gradient function of its type frequency and reliability. Second-order schemas are also notoriously difficult for learners to acquire (e.g., Braine et al., 1990, cf. Audring, 2019, p.14). Learning productive second-order schemas appears to require either encountering the corresponding schemas in the same context, where one form is expected but the other occurs instead (e.g., Onnis et al., 2008), or encountering them in close temporal proximity, so that the form of one can be used to predict the form of the other (Smolek, 2019). Constructions, as form-meaning mappings appear to be easier to acquire (e.g., Braine et al., 1990; Kapatsinski, 2013).

Given the existence and easy learnability of constructions, it is reasonable to assume that second-order generalizations are learned and used only for patterns that cannot be captured with direct form-meaning mappings (Bybee, 2001; but cf. Booij and Audring, 2017a). Therefore, in order to convincingly argue for the necessity of a second-order schema, we need to show that one could not have understood or generated each of the constructions it links without reference to the other construction. This is a high bar to clear in syntax. For example, hearing *The dax fribbles a wug to the frumbly swuppet*, the listener does not need to activate the alternative double-object formulation to understand the sentence (Goldberg, 2002). S/he also does not need to use this specific formulation in generating the double-object alternative. Hearing the sentence, the listener could categorize the *swuppet* as an animal or human and perhaps assign it a gender, necessary for choosing the pronoun. When producing the sentence, the speaker would then be influenced by the inferred characteristics of the *swuppet*, the *wug*, and the action of *fribbling*, which do not require reference to the prepositional dative formulation—they are inherent to the inferred semantics. Choosing to use a pronoun requires knowing that the *wug* and/or the *swuppet* was mentioned, but also does not require reference to the prepositional dative formulation. The choice of the construction depends on this choice—the double object construction is strongly favored by selecting a pronoun to refer

to the swuppet (Bresnan et al., 2007) and disfavored by a long noun phrase—but does not depend on anything about the double object formulation. In other words, the two constructions are not in a feeding relationship—generating one of these constructions does not require reference to anything that one could find only in a construct of the other. The closest one comes to such a relationship is when the other construction would sound awkward given a certain filler for the first NP because of that filler’s phonology (*?I gave the highly agitated swuppet that was zwigging all over the room a wug*; see Shih, 2017, for a recent review). However, even such cases do not require the use of a second-order schema. For example, the speaker could begin to generate both formulations in parallel, and the awkward-sounding construction would simply lose the race because it is harder to formulate.

In contrast to syntax, paradigmatic morphology presents numerous examples where one does need to reference the form of one construction to generate a related one. For example, Becker and Gouskova (2016) documented the productivity of the generalization that...oCC#Nom.Sg~...eCCa#Gen.Sg but...CoC#Nom.Sg~...CCa#Gen.Sg in Russian. Here, the same form would result from vowel deletion in the Genitive in both cases (...CC#), but it is avoided when *another* form of the word ends in a single consonant (...oC#). Thus, generating the Genitive Singular seems to require reference to the Nominative Singular construction. It is difficult to find any comparable examples of syntactic constructions; that is, constructions whose use or form depends in an arbitrary fashion on the form of another construction.

Second-Order Schemas Help Enact Large Changes to the Base

Second-order schemas allow the speaker to enact arbitrary changes to an activated form when constructing a production plan. Evidence for this claim comes from a recent dissertation by Smolek (2019), who exposed participants to a language with second-order schemas and manipulated how easy they were to extract from the input. She then tested speakers’ knowledge of the language using both judgment and production. She found that participants would produce large changes to the base only if second-order schemas were easy to notice in training. Acceptability judgments were unaffected, as was production of smaller, and more *a priori* likely changes.

In Smolek (2019), a subset of singulars mapped onto plurals ending in [tʃa], undergoing a stem change either when they ended in [k] or when they ended in [p], as shown in (1), where the consonants in curly brackets were presented to different participants:

- (1) ... *blu{k;p}SG blutʃaPL smi{p;k}SG smi{p;k}aPL klatsG
klataPL...*

Learners produced the $p \rightarrow tʃ$ change only if exemplifying singular-plural pairs were kept intact during training as in (1-2). When only faithful pairs were kept intact (3) or all words were presented in random order (4), participants did not learn to produce the stem change, retaining the [p] of the singular.

- (2) ... *smikSG blupSG blutʃaPL klataPL smikaPL klatsG...*
(3) ... *blupSG klatsG klataPL smikSG smikaPL blutʃaPL...*
(4) ... *smikSG blupSG klataPL smikaPL klatsG blutʃaPL...*

When exposed to $p \rightarrow tʃ$ using the random order in (4), participants judged singular-plural pairs exemplifying the stem change as being more acceptable than those without the change. In both judgment and production, they also did not know what stems should change and what stems should not, indicating that they had not learned paradigmatic, second-order mappings. However, they would not change any stems while judging that all stems *should* change. Because this was not true of the smaller change $k \rightarrow tʃ$, where judgments and production probabilities aligned, Smolek argued that second-order schemas are particularly important for making large changes. Without a second-order schema, one can still judge unexpectedly frequent constructions like $tʃa \sim PL$ as being particularly characteristic of the experienced language but would not produce such outcomes from inputs that are either very different or *a priori* unlikely to map onto them.

Smolek’s results are partially consistent with Booij and Audring’s (2017a) proposal that “output-oriented, constructional schemas [i.e., form-meaning mappings that do not make reference to other forms] should be used for stating regularities that are not productive” because “these schemas have a motivational function only” (p.59). However, I would not go that far, as there is evidence that constructional schemas can be productive and can even be used in preference to second-order schemas with which they conflict (e.g., Wang and Derwing, 1994; Kapatsinski, 2013). Furthermore, in Smolek’s (2019) experiments, constructional schemas could support productive generation of plurals, except when singular-plural mappings in the input involved large unexpected changes to the base.

Second-Order Schemas vs. Rules

Second-order schemas are typically depicted as bidirectional (e.g., Booij, 2010; Jackendoff and Audring, 2016), like first-order schemas/constructions mapping form and meaning. However, this appears to be a false analogy. Unlike forms and meanings, paradigmatically-linked words do not occur at the same exact time in one’s experience—one or the other word occurs first and can then be used to predict the other. Being able to produce plurals from singulars also does not guarantee being able to produce singulars from plurals, and depends on the reliability of the mappings *in that particular direction* (e.g., Engelmann et al., 2019). Thus, paradigmatic mappings linking two forms often have different strengths in different directions.

As directed paradigmatic mappings, second-order schemas resemble rules that map surface forms onto other surface forms (as proposed by Albright and Hayes, 2003). They differ from such rules because they do not involve a split into a change and a context in which that change occurs (Kapatsinski, 2012, 2013; Jackendoff and Audring, 2016). A rule is an operation that occurs in a certain context; a second-order schema is a mapping between two constructions (or their forms). I have argued for schemas over rules by observing that mappings that involve different changes but the same output can “conspire”: as evidence for

one increases the other becomes more productive alongside it, and participants who like or frequently produce one mapping also tend to like and frequently produce the other. In particular, adding pairs of words exemplifying [...tʃ]_{SG}~[...]PL to a language in which [...k]_{SG}~[...]PL but [...]SG~[...]PL led participants to overgeneralize the k→tʃ change to [t] (Kapatsinski, 2012, 2013). These results suggest that learners are treating [...tʃ]_{SG}~[...]PL and [...]SG~[...]PL as exemplifying the same schema even though they involve different changes (0→i vs. t→tʃi). This result rules out models such as Albright and Hayes (2003) or Becker and Gouskova (2016) that split words into changes and contexts, the ingredients of a rule, and then generalize only over the contexts in which a particular change occurs.

However, rules can be rescued if we assume that zero is not a possible (or likely) input to the change, so any change must involve at least one overt segment as the input. That is, learners presented with examples like *blutʃ*_{SG}~*blutʃi*_{PL} are experiencing the change tʃ →tʃi rather than 0→i. The results of Kapatsinski (2012, 2013) are then captured by assuming that learners do generalize over changes, and that they generalize over inputs more than over outputs so that all kinds of inputs initially map onto [tʃi]. Assuming that outputs are action plans to be performed, greater generalization over inputs than over outputs may be a general property of learning in a world where cues calling for a certain action can vary but actions need to be performed with some precision to be efficacious (Kapatsinski, 2018, pp.64-66).

Another way to test the difference between mappings and operations is afforded by subtraction. Pure subtraction involves removing a fixed unit regardless of what remains, in contrast with truncation, which refers to removing as much material as necessary to fit a fixed prosodic template. Truncation is easily captured by a construction in which the form has certain prosodic characteristics. Inkelas (2015) identifies a diachronic pathway from subtraction to truncation, which suggests that speakers often extract a construction from the truncated forms produced by subtraction. However, subtraction does appear to be learnable, and is not easily captured by constructions.

Learnability of subtraction was examined in Kapatsinski (2017a; 2018, pp.186-192). Native English speakers were exposed to artificial languages that could be interpreted as exemplifying either truncation or subtraction. In these languages, the final vowel of the singular was deleted to form the plural always resulting in CVCVC. These languages then could be interpreted either as using the construction [CVCVC]_{PL} or the rule / operation V→0/[_]PL. At test, participants were then presented with CVCV singulars, for which the two generalizations predict different choices: satisfying the construction would involve the operation of consonant addition (unattested in training), but result in an attested product (CVCVC), while following the rule would involve using the attested operation of vowel deletion to produce an unattested product, CVC. Participant choices depended on whether one of the consonants was overrepresented at the ends of singulars. Participants were more likely to add a consonant if they knew which consonant to add (the overrepresented one). However, both strategies were

attested, sometimes within the same individual. These results therefore seem to provide support for both constructions and rules. In the next section, I explore two ways of capturing subtraction within a constructionist framework, without the use of rules.

Subtraction Without Rules: Conditioned Copying or Negative Associations

Subtraction is difficult to capture with a second-order schema because it involves mapping something onto nothing, and null elements are not part of the constructionist framework. How then can subtraction be incorporated into the constructionist worldview? In Kapatsinski (2017b; 2018, pp.193-199), I argue that constructions must be supplemented by an operation that there I called *copying*, on analogy with the copy connections of recurrent networks (Elman, 1990). To produce anything, the speaker needs to construct and execute a production plan, and constructions stored in long-term memory compete for being incorporated (“copied”) into the plan.

Subtraction involves learning *not* to copy a certain element of an activated form when expressing a certain meaning. Thus, it may be captured by making copying conditional on various aspects of the input (Kapatsinski, 2017b). Thus, if we assume that copy connections are gated, these gates may be closed by certain meanings and input forms. If production plans for wordforms are filled out left-to-right (Roelofs, 1999), then it may be sufficient for alternative segments to be competing for a “future” slot in the plan. Preventing copying of a final vowel into plurals would then involve learning a negative weight for a connection from the semantics of plurality to a gate on the copy connection that would make the final vowel the future: $w(\text{PL} \rightarrow [\text{V}\# \rightarrow \text{future}]) < 0$.

Because copying of activation patterns from one brain area to another is biologically implausible (Grossberg, 1987), the construction of a production plan is likely implemented as establishing a resonance between parts of a control structure (e.g., “future”) and activated form units. However, it is not clear how the formation of a resonance can be conditioned, thus it is worth considering alternatives to conditioned copying. A possible avenue to accounting for both changes to the base and subtraction is to incorporate negative meaning-form associations. In current constructionist frameworks, construction forms are templates that are filled out by material from long-term memory (see Jackendoff and Audring, 2016). As such, they can only be positively associated with the meaning they express and lack a mechanism for capturing subtraction when it is independent of the resulting shape of the product. However, any computational model that learns associations between meanings and the forms that cue them also learns negative associations between forms and meanings (as stressed by Ramscar et al., 2014, and Roembke et al., 2016).

Negative form-meaning associations can account for subtraction. For example, the final vowel deletion pattern in Kapatsinski (2017a) could be described as a negative association between PL and V#. They can also account for stem changes, the meaning to be expressed inhibiting elements of the base that are dissociated from it. For example, to produce a singular from a

known plural, an English speaker would inhibit the -s suffix via a negative SG→s# association. The existence of such an association receives independent support from the fact that the singular form *lens* is often misspelled as *lense*, often enough for both to be entered in dictionaries. This misspelling is motivated by the fact that an s#, and especially a Cs# indicates plurality. The intention to produce an adjective may inhibit nominalizers that distinguish adjectives from nouns (see the next section for an example). If negative associations are particularly strong for unexpectedly absent elements of form, this account may also account for stem change examples like k→s in English -ity nouns. A learner of English would expect a [k] after *electrici...* Not hearing it would provide evidence that it is suppressed by the meaning the speaker was expressing. The element of the meaning that discriminates *electricity* from *electric* is whatever distinguishes nouns from adjectives. That element of meaning would then activate -city and suppress [k]. Similarly, it is possible that the Genitive Singular in Russian inhibits oC# as well as activating CC#, resulting in greater deletion of vowels from...oC# singulars than from...oCC# singulars in Becker and Gouskova (2016). Having heard the frequent Nominative/Accusative *puʃok* "little furball," the speaker would expect the same form in a subsequent production of the rarer Genitive; hearing *puʃka*, s/he would then learn that the Genitive Singular disfavors the oC# as well as favoring the -a#.

Cases that would still require second-order schemas involve patterns in which the same structure can be either favored or disfavored in a certain paradigm cell depending on the corresponding form in some other cell. For example, in deriving a Russian Genitive Plural from a known Nominative Singular, /o/ can be both deleted and inserted, depending on whether the noun is Masculine or Feminine, a difference that can be predicted from Nominative Singular forms: *mis-k-a*→*mis-ok* "bowl" but *kus-ok*→*kus-k-ov* "piece." Here...ok# appears to be both eliminated by the Genitive Plural (for Masculines) and imposed by it (for Feminines). Unless these types of choices can be attributed entirely to semantic differences between the word classes (in this case Masculines and Feminines), they require productive second-order schemas. Interestingly, the Genitive Plural is exactly the paradigm cell that Russian speakers have a difficult time filling; with great uncertainty regarding the correct form. For example, dictionaries record both *portkov* and *portok* as the Genitive of the pluralia tantum *portki* "pants," which could be either Masculine or Feminine as unambiguous forms are missing. Paradigm gaps abound, and are spreading (Daland et al., 2007). For example, there is no Genitive Plural for *metʃta* "dream" even though there is one for *matʃta* "mast." The difficulties make sense if the production of this form relies on second-order generalizations, since such generalizations are difficult to acquire.

WHAT GOES INTO ONE MORPHOLOGICAL CONSTRUCTION

From this perspective, a production plan for a novel word is a blend of a number of units stored in long-term memory and activated in parallel by the intended meaning. This results

in forms being multiply motivated (Taylor, 2012; Kapatsinski, 2013; see also Burzio, 1998; Booij and Audring, 2017b). As an example, consider the [*bez...*]_A construction in Russian, exemplified below. This construction carries the same meaning as the [*...less*]_A construction in English (*groundless*, *priceless*). In Russian, the prefix is the same form as the preposition *bez*, "without" and has grammaticalized out of it. I have collected all 341 examples of this construction from the 125,000-word reverse dictionary of Russian (Zaliznjak, 1974).

I will argue that this construction represents a blend of prepositional phrases of the type [*bez* N.GEN] and adjectives, as well as properties associated with the to-be-expressed meaning. For example, in (5)-(6), the adjective "costless" (i.e., free) is motivated by both the prepositional phrase "without cost," which shares *bez* with the adjective, and the adjective "costful" (i.e., not free). In particular, it contains the adjectival suffix -n, which is shared with the *bez*-less adjective. Whenever a *bez* adjective has a *bezless* counterpart, the two share the adjectivizing suffix. However, 31% of *bez* Adjectives lack a counterpart without *bez*, exemplified in (7) and (8). For these adjectives, the only possible base is the corresponding prepositional phrase. However, there are also (less numerous) examples in which the *bez* adjective has no corresponding prepositional phrase, shown below in (18)-(21).

- | | | | |
|-----|--|-----------------------------|---------------------------|
| (5) | [<i>bez</i> - <i>plat</i>]-n-yj | | <i>bez</i> <i>plat</i> -y |
| | without-cost-A-MASC.SG.NOM | | without cost-SG.NOM. |
| (6) | <i>bez</i> -[<i>plat</i> -n-yj] | | <i>plat</i> -n-yj |
| | without-cost-A-MASC.SG.NOM ³ | | cost-A-MASC.SG.NOM |
| (7) | <i>bez</i> - <i>kryl</i> -yj | <i>bez</i> <i>kryl</i> -jev | <i>kryl</i> -at-yj |
| | "wingless" | "without wings" | "winged" |
| (8) | <i>bez</i> - <i>lik</i> -yj ⁴ | <i>bez</i> <i>lik</i> -a | ?? |
| | "faceless" | "without face" | |

Choosing the Suffixes: Schematic and Syntagmatic Conditioning

The final suffix is the case-gender-number agreement marker and is almost regularly -yj in the dictionary (Masculine Singular Nominative) form, with the exception of *bez-mater-n-ij* "motherless," which likely avoids homophony with *bez-mater-n-yj* "lacking taboo words," *bez-mu3-n-ij* "husbandless," which follows the same pattern, and *bez-trud-ov-oj* "laborless," which shares the -ov-oj with its much more frequent *bez*-less pair *trud-ov-oj* "labor-A." With the exception of *trud-ov-oj*, *bez*-less adjectives ending in -oj correspond to *bez*- adjectives ending in -yj (e.g., *vyez-d-n-oj* "able to leave" but *bez-vyez-d-n-yj* "unable to leave," *tsvet-n-oj* "colorful" but *bez-*tsvet-n-yj** "colorless"). The

³The Masculine Singular Nominative is the one form in which more than one inflectional suffix is possible in adjectives. In all other forms, adjectives inflect regularly.

⁴Phonetically, this word ends in is [kʲij] or [kʲj] depending on dialect but this is due to language-wide phonotactic constraints and is not conditioned by anything specific to the construction thus I am abstracting away from it here. This choice should not be taken as an endorsement of phonemes or underlying forms as a psycholinguistic construct.

choice of suffix also comes with a choice of stress location, as *-yj* is unstressed, while *-oj* is stressed. While the number of such pairs is low ($n = 4$), they suggest that *-yj* must be activated by the meaning of the construction.

At the same time, *-yj* must also be strongly associated syntagmatically with the preceding adjectivizing suffix *-n*, as 95.5% of *-n* adjectives take *-yj*, with only 3.5% taking *-oj* and 1% *-ij*. Compare the very low rate of *-oj* use after *-n* to its rate of use after another adjectivizing suffix, *-ov*. While *-yj* is still dominant with this suffix, accounting for 74% of the adjectives, *-oj* accounts for 26%, which is significantly higher than the 3.5% seen with *-n* ($p < 0.00001$ by Fisher exact test). As noted earlier, the only instance of *-oj* use with *bez-* occurs after the suffix *-ov*. These results suggest that there are syntagmatic associations between *-ov* and *-oj*, and between *-n* and *-yj*, even though the “A” meaning generally is associated with *-yj* more strongly than with *-oj*.

The adjectivizing suffix is not fully predictable. However, 79% of *bez-* adjectives bear *-n*, and *-n* also accounts for the majority of *bez* adjectives that lack *bezless* counterparts (67%), i.e., pairless adjectives. This is a significantly higher percentage than for adjectives generally, where *-n* accounts for ~52% of types ($p < 0.00001$ by Fisher exact test). Therefore, *-n* may be considered to be part of the construction, activated by its meaning (“WITHOUT N”)⁵.

There are also many pairless adjectives without an adjectivizing suffix, as in (10)-(11). These form 23% of pairless *bez-* adjectives while only one suffixless *bez* adjective, *bez-pal-yj* “fingerless” has a *bez-*less pair in the dictionary ($p < 0.00001$), and that pair is now obsolete. Suffixless formations are semantically conditioned: all adjectives referring to lack of expected body parts are formed this way; animal body parts account for 20/25 such adjectives. The remaining adjectives refer to parts of non-animal “bodies,” formed from the roots *verx* “top,” *list* “leaf,” *os'* “axle,” and metonymic extensions, *pol* (“sex/gender”) and *styd* (“shame”). Interestingly, the body part semantics cause a suffixless formation only if the body part is in some notable state: thus, *bez-kryl-yj* “wingless” and *širok-o-kryl-yj* “wide-winged” but *kryl-at-yj* “winged”; *bez-puz-yj* “belly-less,” *tolst-o-puz-yj* “fat-bellied” but *puz-at-yj* “bellied”; *bez-golos-yj* “having no voice” and *gromk-o-golos-yj* “having a loud voice,” but *golos-ist-yj* “having a [good] voice.” The adjectivizing suffixes that are removed from such adjectives in forming the *bez-* form are *-at*, *-ist* and *-ast*. They must be suppressed by the “remarkable state of a body part” semantics.

The suffix *-(l)iv* is always shared with the *bez-*less adjective and thus not associated with “without.” Its selection is independently semantically conditioned, in that it refers to characteristic behaviors/character traits. Thus, an *o-pas-liv-yj* “cautious” person operates with caution (*opas-k-a*), and a *za-stentš-iv-yj* “shy” person lives behind a self-imposed wall (*stenka*), having the quality of with *za-stentš-iv-os'š* (“shyness”). A *zabot-liv-yj* “caring” person performs *zabot-a* “care” for other people.

⁵The suffix has a relatively rare allomorph, *-enn*, that attaches to stems ending in certain consonant clusters (*stv*, *{d,t}r*, *zn*) but not *{z; 3}d*, *{s;k;r;n}t*, *ltš*, or *lk*. Thus *-enn* occurs where there would be a sonority sequencing violation if *-n* were attached. Because apparent allomorphs often have additional semantic conditioning (Endresen, 2015), *-enn* tokens are not included in the *-n* counts above.

The choice of variant is syntagmatically-conditioned: *-liv* after coronals, */s/* and */t/*, and *-iv* after */tʃ/*.

Copying From the Prepositional Phrase

The suffix *-ov* is less common in *bez-* adjectives than in other adjectives and must therefore be inhibited by the construction's meaning: only 5 (1%) of *bez-* adjectives have the suffix compared to 14% of all adjectives ($p < 0.00001$). Interestingly, this suffix is of the same shape as the Genitive Plural Masculine inflectional suffix on nouns (*kot* “tomcat,” *kot-ov* “tomcats-GEN.PL.MASC”). The preposition *bez* requires a Genitive noun, but does not place requirements on number or gender. Thirty-one percentage of the nouns in PP's corresponding to *bez-*Adjectives take *-ov* in the Genitive Plural. However, all nouns corresponding to *bez-* Adjectives taking the suffix *-ov* bear the Genitive Plural suffix *-ov*. While there are only five such adjectives, the pattern is suggestive of *-ov* being copied from the noun in the corresponding prepositional phrase. The pattern is statistically robust across the class of *-ov* adjectives in the dictionary where 74% (1022/1373) have a corresponding noun ending in *-ov*, a proportion statistically greater than the 39% observed with *-n* adjectives ($p < 0.0001$). Thus, it appears that the adjectival suffix *-ov* often results from a genitive plural noun inflection copied into the adjective when the adjective is formed. Copying of inflectional suffixes into adjectives, where they look like derivational, adjectivizing suffixes suggests that copying operates on a fully inflected wordform rather than a stem, and that what is being copied are surface chunks from that form. At the same time *-ov* cannot *always* result from nominal inflection because not all such adjectives have nominal bases ending in *-ov*. In 26% of the cases, it is imposed directly by the A meaning.

Copying from the prepositional phrase is also supported by another aspect of the forms of *bez-* adjectives, the spelling of *bez-* (Kapatsinski, 2010b). Both the prefix and the preposition undergo voicing assimilation, so that *bez* is pronounced [bes] before voiceless obstruents. However, the spelling rules for the prefixes differ from those for prepositions: the preposition must always be spelled *bez*, whereas the prefix must be spelled the way it sounds, with *<s>* before voiceless obstruents. Kapatsinski (2010b) shows that Russian college students spell the prefix [bes] *<bez>* ~50% of the time in low-frequency *bez-* adjectives they do not know, even in a graded dictation test. The error rate is two orders of magnitude higher than the error rate for other comparable prefixes (*iz-* and *raz-*), which also end in */z/* and obey the same spelling rules. Like *bez-*, the errorless prefixes have homophonous free morphemes that are always spelled with *<z>*. In the case of *iz-*, as in the case of the error-prone *bez-*, the free morpheme is a preposition i.e., near-synonymous with the prefix and is the diachronic source of it. However, neither *iz-* or *raz-* verbs have bases that contain free morphemes that correspond to the prefix and from which its spelling can be copied. Both prefixes derive perfective verbs from imperfective ones as in (9)-(10). The low rate of spelling errors on *iz-* and *raz-* suggests that the spelling errors on *bez-* are due to writers copying the orthographic *<z>* of the base prepositional phrase into the production plan for the adjective. Frequent *bez-* adjectives are spelled correctly because their orthographic forms can be retrieved from the lexicon.

- (9) *raz*_{<s>}-*kop-a-tʃ* *kop-a-tʃ*
 apart-dig-V-INF dig-V-INF
*raz*_{<z>} *kop-a-tʃ*, to...
 since [we are to] dig-V-INF, then
- (10) *iz*_{<s>}-*kop-a-tʃ* *kop-a-tʃ* *iz*_{<z>} *korobk-i*
 out-dig-V-INF dig-V-INF out-of box-GEN.SG

No Single Base Is Necessary

A parallel, associationist construction predicts that there should be no single base from which *bez* adjectives are derived (see also Burzio, 1998; Booij and Audring, 2017b). The forms blended into the plan need to meet only one criterion: they need to be associated with, and therefore recurrently activated by the intended semantics. The more strongly a form is activated, the more it is predicted to affect the blend. This hypothesis is strongly motivated by results on the diachronic phenomenon of paradigm leveling, which happens between forms that are strongly related semantically (Bybee and Brewer, 1980), and changes less frequent forms by blending in elements of more frequent ones (Tiersma, 1982). For example, Bybee and Brewer (1980) show that paradigm leveling in Provençal verbal paradigms happened between forms of the verb that share all inherent semantics, differing only in agreement. Tiersma (1982) provided evidence that Frisian nouns have leveled mostly in favor of singular forms, except for those for which the plural form is more frequent.

Note that, in any case of paradigm leveling, there is a form that would fully express the intended semantics. This form would receive more activation from the semantics if frequency were controlled, and therefore can often prevent other forms from affecting the blend, blocking/pre-empting the formation of synonyms. Leveling occurs when the form fails to block the formation of a synonym because it is not accessible enough from the intended meaning, and is replaced by something else. That something else is, furthermore, not another existing form, but a new formation that incorporates elements of the more frequent semantically similar form into the form that matched the intended semantics fully. The existence of this process strongly implicates parallel activation of competing forms and a blending process that can combine them into a novel production plan. In the case of *bez*- adjectives, semantic similarity explains copying from the corresponding prepositional phrases that can express most of the meaning of the *bez*- adjective. Because these phrases contain Genitive nouns, this also explains why it is the Genitive i.e., copied.

The proposal that words are formed by blending forms activated in parallel by the intended meaning contrasts with the hypothesis that there is a single base for any particular type of morphologically complex word (Albright, 2002). We have already seen evidence that *bez* adjectives are motivated by both *bez*-less adjectives and prepositional phrases, contradicting the single base hypothesis. However, until now we could maintain that there is always a prepositional phrase base, suggesting that we could claim that there is one particular base i.e., necessary for deriving a *bez* adjective. However, the problems go deeper: first, it is not possible to claim that the forms of the nouns in the base prepositional phrases always come from the same paradigm cell;

second, there are *bez* adjectives that do not have corresponding prepositional phrases.

Some Russian nouns have different stem forms in Singular and Plural Genitives. The examples in (11)-(13) show that it is not: some adjectives copy the plural form (11) while others copy the singular (12). Sometimes, different *bez* adjectives can even be derived from the forms in different paradigm cells, as in (12)-(13). Thus, a single base paradigm cell cannot be identified: whatever forms match the intended semantics best are the ones copied.

	Singular	Plural
(11) <i>bez-det-n-yj</i> childless	<i>bez diʃ-a/rebʲonk-a</i> without child- GEN.SG	<i>bez det-ej</i> without children- GEN.PL
(12) <i>bez-tʃelovek-tʃ-n-yj</i> inhumane	<i>bez tʃelovek-a</i> without person- GEN.SG	
(13) <i>bez-ʃud-n-yj</i> empty of people	? <i>bez ʃud-a</i> ⁶ without folk- GEN.SG	<i>bez ʃud-ej</i> without people- GEN.PL

A single base is also ruled out by the fact that the base noun can lack an acceptable Genitive Plural form (the Genitive Plural is the nominal form in which paradigm gaps are common in Russian), be uncountable and thus lack plural forms altogether or, conversely, be a pluralia tantum that lacks singular forms. In such cases, the available form of the noun must be used to produce the adjective. For example, *bez-vred-n-yj* “harmless” cannot be formed from a plural form of *vred* “harm” because it is not countable and lacks plural forms. Conversely, *bez-ʃtan-n-yj* “pants-less” must be formed from the plural (*ʃtan-ov*) because it lacks a singular form.

There are also cases of variation, as in (14). Note that retention of the Vn is consistent with the adjectives being motivated by prepositional phrases, as it is not present in the singular Nominative or Accusative forms but is present in the Genitives requires by *bez*:

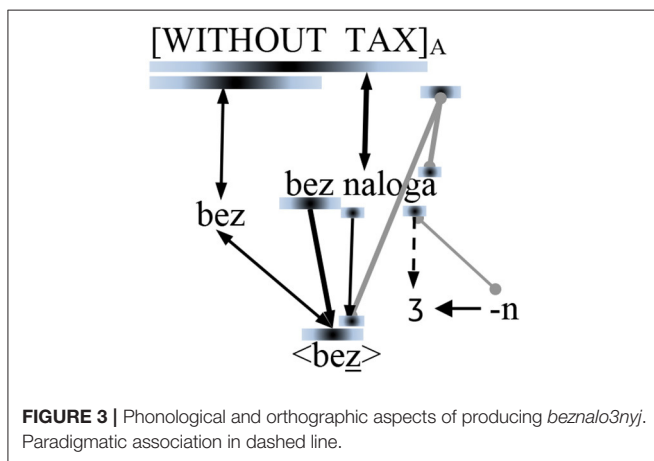
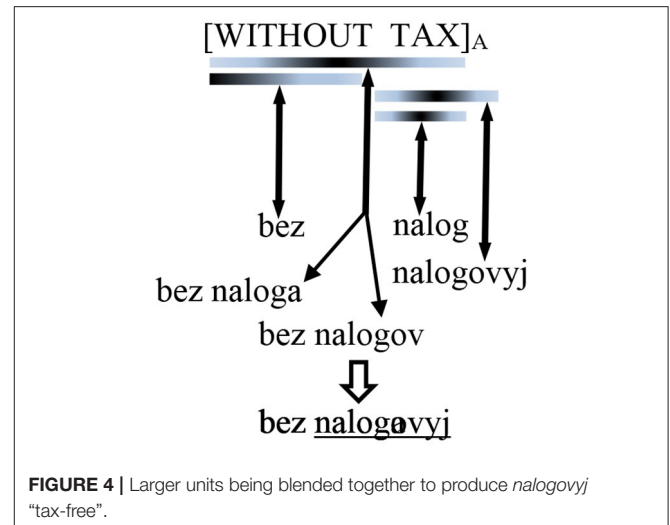
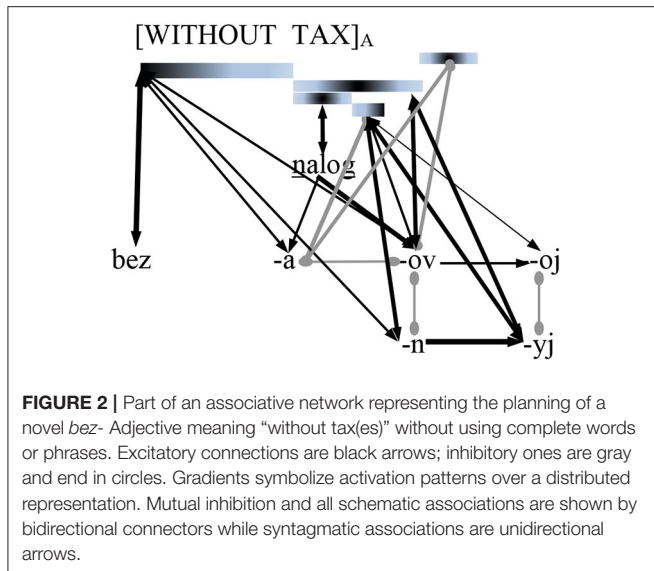
- (14) *bez-sem{e;ja}n-n-yj* *bez semen-i*
 seedless without seed-GEN.SG
bez semjan
 without seeds-GEN.PL

While the vast majority of *bez*- adjectives have a corresponding prepositional phrase, some do not, indicating that *bez*- adjectives cannot always be derived from prepositional phrase bases. Thus, the adjective in (15) appears to be formed directly from a verb.

- (15) *bez-voz-bran-n-yj* **voz-bran-ije voz-bran-iʃ*
 “unchallenged” “prohibit”

Other adjectives formed from verbs can often be identified because they retain the infinitival inflection from the base verb, and add the sequence *-elʃ-n-yj* (16-18). The *-elʃ* is the agentive marker (cognate with English *-er*), as in *stroitʃ* “to build” ~

⁶This is a collective noun referring to a “type” of people and is awkward without an adjective defining the type, such as “city” or “working.” The plural of this example is the only plural for *tʃelovek*.



- (16) *bez-nrav-stv-enn-yj* **nprav-stv-o* *nprav-itʹ-sa*
 “moral” “to be liked”
bez nprav-stv-enn-ostʹ-i
 “without morality”
- (17) *bez-ot-lag-at-elʹ-n-yj* **ot-lag-atʹ-elʹ* *ot-lag-atʹ*
 “urgent” “to delay”
bez ot-lag-atʹ-elʹ-stv-a
 “without (purposeful) delay”
- (18) *bez-ot-nos-itʹ-elʹ-n-yj* **ot-nos-itʹ-elʹ* *ot-nos-itʹ-sa*
 “unrelated” “to relate”
 ??*bez ot-nos-itʹ-elʹ-n-ostʹ-i*
 “without relativity”

Producing a bez-Adjective

This section provides an informal illustration of what production looks like in this framework⁷. The example shows the process of generating a novel adjective the meaning “tax-free,” an adjectival equivalent of “without tax(es).” This adjective is not in the dictionary but can be found on the web, with the two possible forms *beznalogovyj* and *beznalo3nyj*. The former is much more common, with 418 vs. 56 Google hits, and intuitively appears more acceptable. I take the grammar to be responsible for generating both forms and explaining why the former is more common. **Figure 2** shows some of the schematic and syntagmatic associations involved in producing a novel *bez* adjective, including only morphemic chunks. **Figure 3** illustrates the role of paradigmatic associations in enacting changes to the base, and the role of the base in the orthography, showing that chunks larger than the morpheme also play a role in production. **Figure 4** illustrates how blending of these larger forms would result in the most common form produced. Note that **Figures 2, 4** should not be seen as two different “routes” for producing the new adjective: there is instead a near-infinite number of

stroitʹelʹ “builder.” However, these adjectives are not derived from such nouns: the nouns are often missing, and retaining the semantics of the *-er* would require adding a different adjectivizing suffix, thus *stroitʹelʹ-n-yj musor* “building garbage” (i.e., garbage associated with building something), vs. *stroitʹelʹ-sk-yj musor* “builder garbage” (garbage associated with a builder or builders generally). The adjectives can usually be related to “deadjectival” nouns ending in *-stv-o* or *-ostʹ-i* (*stroitʹelʹstvo* “the process of building”). The examples in (16) and (17) are difficult to explain without reference to such a noun. However, the example in (18) is difficult to relate to the corresponding noun: the corresponding PP is awkward and not interpretable as synonymous with the adjective. Once again, *bez-* adjective forms are produced using whatever semantically close words are available, as one would expect from a lexicon i.e., structured as a parallel, associative network.

⁷A formal treatment would spell out associations as weighted constraints (e.g., Boersma, 1998; Burzio, 1998; Kapatsinski, 2013) but this is beyond the scope of this paper.

routes because all meaning-to-form associations activated by as semantic feature are activated in parallel.

The top of the diagram in **Figure 2**, [WITHOUT TAX]_A represent the intended meaning, which I assume to be a distributed representation, as symbolized by the gradients below it. The top gradient represents the unique aspects of the meaning of the *bez*-A construction, which distinguish that construction from all other constructions and make its representation more than the sum of its parts. The next gradient down is the meaning “without,” which strongly cues and is cued by *bez*, as shown by the thick bidirectional arrow. The next one down is the meaning of “relating to taxes,” or [tax]_A, for which there is an established adjective, *nalogovyy*. The gradients for “tax” and the Adjective category follow.

The meaning “without” is consistent with both adjectives and prepositional phrases and therefore activates both Genitive case markers appropriate for the prepositional phrase and the adjectivizing suffix *-n* i.e., favored over others by this construction. The Genitive suffixes activated include *-a* and *-ov* appropriate for a noun like *nalog* and suffixes from other declension classes (not shown here). The activation of *nalog* from the meaning “tax” boosts the Genitive suffixes appropriate to its over their competitors from other declensions. This is shown by the arrows from *nalog* to the two suffixes *-a* and *-ov*. Because the two suffixes are incompatible with each other, I assume an inhibitory connection between them.

We have seen evidence that the suffix *-ov* occurs in *bez* Adjectives primarily when the corresponding noun selects it as a Genitive Plural suffix. Thus, *-ov* in prepositional phrases and adjectives with *bez* appears to be the same form, associated with the meaning “without,” which is the meaning of *bez* and one meaning of the Genitive. Because *-ov* can serve as an adjectivizing suffix, it must also be activated by the Adjective category. Interestingly, however, *-ov* is disfavored by *bez*-Adjectives compared to other adjectives. It must therefore be inhibited by the meaning of the construction as a whole even though it is favored by all of the parts of that meaning (“without,” “tax,” and “A”). **Figure 2** therefore includes an inhibitory connection from the top gradient (unique features of the construction) to *-ov*.

The adjectivizing suffixes are syntagmatically associated with the case-number suffixes that follow them. As shown above, *-yj* is more common than *-oj* across the board but is particularly rare after *-n*. For this reason, the top-down connection from A to *-yj* is stronger than the one to *-oj* and *-oj* is syntagmatically boosted by *-ov* while *-yj* is boosted by *-n*. In addition, *-ov* and *-yj* are both activated by the “[tax]_A” meaning because *nalogovyy* is an existing adjective.

Figure 2 predicts rapid activation of *bez*- and *nalog*, which are not inhibited by anything. At this point, the speaker’s intended production is the same whether or not it resolves into a prepositional phrase or an adjective, because both constructions are compatible with most of the meaning intended. This partial overlap results in competition between the two constructions in usage. According to **Figure 2**, which construction ends up being produced depends on resolution of two competitions: between *-a*, *-ov*, and *-n* and between *-oj* and *-yj*. The first competition

will resolve mostly in favor of *-n* because the intended meaning inhibits the other competitors. The second competition will be resolved in favor of *-yj*, which receives more activation from the intended meaning and from the preceding element, and is also favored by the more likely preceding element. Because all processing happens in parallel, it is possible for the competition between *-yj* and *-oj* to resolve before the competition between *-ov* and *-n*, in which case *-yj* is expected to help select *-n* using a backward syntagmatic association (not shown), rather than *-n* helping select *-yj*.

Figure 3 shows additional aspects of form generation, specifying phonology and orthography. Because the suffix *-n* does not allow a velar to precede it, it inhibits the final [g] of *naloga* and *nalogov* if selected, and activates [3], alongside other consonants. The specific consonant, [3], however, is selected because it is also activated paradigmatically by the [g] of *nalog* (dashed arrow). Finally, the orthographic form activated most strongly by *bez* is <bez>, its most common spelling and the only one allowed in prepositional phrases. The strength of this connection could explain why Russian speakers would often spell *bez* with an <s> even when it is a prefix and pronounced with an [s]. However, it does not explain why these errors do not similarly afflict *iz*-, for whom the prepositional spelling is even more common relatively to the prefixal spelling. Thus, the errors must be boosted by the fact that the intended semantics for a *bez* Adjective activate prepositional phrases, while the intended semantics for an *iz* verb do not. This is shown by the connection between *bez* in the prepositional phrase and <bez > in the orthography. Accurate spelling requires the A category to weaken the activation of <z>, allowing the phonological context (here, the [n] of *nalog*) to select the right spelling syntagmatically.

The representation in **Figure 2** therefore oversimplifies the network structure because it omits the larger units like *bez naloga* that are also activated by the intended semantics. Indeed, these units may well be activated by the semantics more strongly than their smaller or less context-bound counterparts: even though smaller units are favored by their greater frequency, larger units match the intended semantics better. This is what allows established forms to outcompete synonymous innovations most of the time. For example, irregular forms like *went* can block/pre-empt the creation of synonymous regulars because *went* is activated by both GO and PAST, whereas each part of *goed* is activated by only one of these cues (Kapatsinski, 2018, p.278). Of course, because frequency also plays a role, blocking can fail, allowing regularization and paradigm leveling to occur.

Figure 4 shows the larger units for the case of *beznalogovyy*. Only units activated by the intended meaning are shown. The block arrow shows that the activated forms are blended into the production plan, by copying and aligning them to maximize overlap. The most likely production, *beznalogovyy*, is predicted. However, blending these larger units will not produce any other variant: the *nalo3nyj* part of *beznalo3nyj* is blocked by the existence of *nalogovyy*. Thus, generating *beznalo3nyj* is possible only using smaller, sublexical units, explaining its lower frequency. Its existence therefore also provides support for the existence of the sublexical route.

DISCUSSION

In this paper, I have argued that constructions are not unitary entities. They emerge from the interaction of schematic (form-meaning), paradigmatic and syntagmatic associations in a parallel, associative network that includes both forms and meanings. Here, I have focused on the role and directionality of schematic and paradigmatic associations and on the proposal that forms are activated in parallel by the intended meaning and blended into a production plan.

I take centrality of symmetrical schematic associations to language production to be a core claim of constructionist approaches (e.g., Bybee, 1985; Goldberg, 2002). There is abundant evidence for the existence of schematic associations and substantial evidence for the assumption that such associations are largely if not always symmetrical. In contrast, paradigmatic associations are likely unidirectional and are of more limited use (cf., Booij and Audring, 2017a). In fact, many isolating languages may get along just fine without paradigmatic mappings. Many native speakers of languages whose description requires arbitrary paradigmatic mappings also do not learn them (Dabrowska and Szczerbiński, 2006; Engelmann et al., 2019). Here, I showed how allowing for negative form-meaning associations further limits the need for paradigmatic mappings. Nonetheless, it is clear that many speakers of languages that require arbitrary phonological mappings between paradigm cells do acquire second-order generalizations, indicating that theories of grammar must allow for their acquisition.

Constructing a new form is a gradual settling process (see Cleeremans, 2004, for a useful simulation), as a “pandemonium” of voices clamoring for or against including various pieces of form into the product being constructed (Kapatsinski, 2013). The resulting form is often a blend of many existing forms. Despite the clamor, the network usually settles on an agreeable solution, although paradigm gaps can emerge when it does not (Albright, 2003)⁸. Generation of new words is a messy and slow process, often taking more than a second, which necessitates the storage of the products for reuse on future occasions, it is also highly flexible, capable of generating an acceptable product by an almost limitless patchwork of routes.

The example of the *bez-* construction illustrates this messy but highly flexible process. Speakers of Russian do generate new *bez-* adjectives as needed—for example, producing *bez-finans-ov-yj* “financing-free” to characterize certain business transactions—by

⁸An important direction for future work is to explain the difference between variation and gaps. That is, why sometimes multiple alternative forms are acceptable, and sometimes none are. Accounting for such cases appears to require distinguishing generation of alternative forms (the focus of this paper) and their evaluation. That is, gaps may arise when all generated forms are subject to a negative evaluation, for whatever reason (social stigma, phonotactics, undesired homonymy, or even aesthetics). Speakers of languages with gaps usually know how the gap could be filled, even though they cringe at the possible fillers.

activating a number of forms that partially fit the meaning to be expressed and blending them together by copying bits and pieces into the production plan. These forms are not always the same forms: whatever forms are available are used. Properties of the construction and the activated chunks of existing forms “clamor” for being copied into the plan. What does get copied depends on how compatible the various chunks are with the meaning to be expressed, on how activated the various base forms are, and perhaps on the speaker’s knowledge of what should and should not be copied.

Some chunks activated as part of existing forms (*-ostʲ* and, less so, *-stv* and *-ov*) will be suppressed by the construction’s meaning, while other chunks may be activated by it directly (chunks like *bez-*, *-n*, *-enn*, and *-yj*, as well as a characteristic pattern of stressing the penultimate vowel). However, the construction’s influence is not absolute. Only some of the meaning to be expressed is part of the “construction”. Semantic features outside of the construction proper such as the fact that the referent lacks a body part may suppress an otherwise dominant *-n* suffix. Frequent forms compatible with the meaning will exert a greater influence than those less frequent and less compatible and may surface in the produced form even if not fully compatible with the construction’s meaning (Bybee and Brewer, 1980; Tiersma, 1982; Harmon and Kapatsinski, 2017).

Often after substantial deliberation, the speaker will settle on a new adjective form with enough confidence to produce it. At that point, the result will be evaluated by the speaker and the interlocutor (e.g., “what a cool way to express that meaning,” “that was hard to pronounce,” or “that was not understood”), stored in their memories (possibly linked to different meanings and divergent evaluations), and will begin its journey through the social and semantic space. As it is reused under circumstances only partially matching the circumstances of its creation, it will be extended to new uses, diffusing away from the speaker and the meaning responsible for its creation (Harmon and Kapatsinski, 2017; Kapatsinski, 2018). Morphology is a mess, and constructions are only a big part of it.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repository and accession number(s) can be found below: <https://osf.io/5er92/>.

AUTHOR CONTRIBUTIONS

VK is responsible for all aspects of this manuscript.

FUNDING

This research was supported by the Faculty Excellence Award from the University of Oregon, to VK.

REFERENCES

- Albright, A. (2002). *The Identification of Bases in Morphological Paradigms*. Doctoral dissertation, University of California, Los Angeles, CA.
- Albright, A. (2003). "A quantitative study of Spanish paradigm gaps," in *Proceedings of the West Coast Conference on Formal Linguistics*, Vol. 22, eds M. Tsujimura and G. Harding (Somerville, MA: Cascadilla Press), 1–14.
- Albright, A., and Hayes, B. (2003). Rules vs. analogy in English past tenses: a computational/experimental study. *Cognition* 90, 119–161. doi: 10.1016/S0010-0277(03)00146-X
- Allen, K., Pereira, F., Botvinick, M., and Goldberg, A. E. (2012). Distinguishing grammatical constructions with fMRI pattern analysis. *Brain Lang.* 123, 174–182. doi: 10.1016/j.bandl.2012.08.005
- Ambridge, B. (2019). Against stored abstractions: a radical exemplar model of language acquisition. *First Lang.* 40, 509–559. doi: 10.31234/osf.io/gy3ah
- Audring, J. (2019). Mothers or sisters? The encoding of morphological knowledge. *Word Struct.* 12, 274–296. doi: 10.3366/word.2019.0150
- Becker, M., and Gouskova, M. (2016). Source-oriented generalizations as grammar inference in Russian vowel deletion. *Linguist. Inquiry* 47, 391–425. doi: 10.1162/ling_a_00217
- Bever, T. G., and Poeppel, D. (2010). Analysis by synthesis: a (re-) emerging program of research for language and vision. *Biolinguistics* 4, 174–200.
- Blevins, J. P., Ackerman, F., Malouf, R., and Ramscar, M. (2016). "Morphology as an adaptive discriminative system," in *Morphological Metatheory*, eds D. Siddiqi and H. Harley (Amsterdam: John Benjamins), 271–302. doi: 10.1075/la.229.10ble
- Boersma, P. (1998). *Functional Phonology*. Doctoral Dissertation, University of Amsterdam
- Booij, G. (2010). *Construction Morphology*. Oxford: Oxford University Press. doi: 10.1093/oxfordhb/9780199695720.013.0010
- Booij, G., and Audring, J. (2017a). "Partial motivation, multiple motivation: the role of output schemas in morphology," in *The Construction of Words: Advances in Construction Morphology*, ed G. Booij (Cham: Springer), 59–80. doi: 10.1007/978-3-319-74394-3_3
- Booij, G., and Audring, J. (2017b). Construction Morphology and the parallel architecture of grammar. *Cogn. Sci.* 41, 277–302. doi: 10.1111/cogs.12323
- Braine, M. D., Brody, R. E., Brooks, P. J., Sudhalter, V., Ross, J. A., Catalano, L., et al. (1990). Exploring language acquisition in children with a miniature artificial language: effects of item and pattern frequency, arbitrary subclasses, and correction. *J. Mem. Lang.* 29, 591–610. doi: 10.1016/0749-596X(90)90054-4
- Bresnan, J., Cueni, A., Nikitina, T., and Baayen, R.H. (2007). "Predicting the dative alternation," in *Cognitive Foundations of Interpretation*, eds G. Bouma, I. Kraemer, and J. Zwarts (Amsterdam: Koninklijke Nederlandse Akademie van Wetenschappen), 69–94.
- Burzio, L. (1998). Multiple correspondence. *Lingua* 104, 79–109. doi: 10.1016/S0024-3841(97)00025-9
- Bybee, J. (1985). *Morphology: A Study of the Relation between Meaning and Form*. Amsterdam: John Benjamins. doi: 10.1075/tsl.9
- Bybee, J. (2001). *Phonology and Language Use*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511612886
- Bybee, J., and Brewer, M. A. (1980). Explanation in morphophonemics: changes in provental and Spanish preterite forms. *Lingua* 52, 201–242. doi: 10.1016/0024-3841(80)90035-2
- Cappelle, B. (2006). Particle placement and the case for 'allostructions'. *Constructions* 2006, 1–28.
- Cleeremans, A. (2004). *Interactive Activation*. Software. Available online at: <https://axc.ulb.be/pages/interactive-activation-application> (accessed November 22, 2020).
- Dabrowska, E., and Szczerbiński, M. (2006). Polish children's productivity with case marking: the role of regularity, type frequency, and phonological diversity. *J. Child Lang.* 33, 559–597. doi: 10.1017/S0305000906007471
- Daland, R., Sims, A. D., and Pierrehumbert, J. (2007). "Much ado about nothing: a social network model of Russian paradigmatic gaps," in *Proceedings of the Annual Meeting of the Association of Computational Linguistics*, Vol. 45, eds A. Zaenen and A. van den Bosch (Prague: Association for Computational Linguistics), 936–943.
- Davis, M., and Redford, M. A. (2019). The emergence of discrete perceptual-motor units in a production model that assumes holistic phonological representations. *Front. Psychol.* 10:2121. doi: 10.3389/fpsyg.2019.02121
- Deacon, T. W. (1997). *The Symbolic Species: The Coevolution of Language and the Brain*. New York, NY: Norton.
- Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychol. Rev.* 93, 283–321. doi: 10.1037/0033-295X.93.3.283
- Diessel, H. (2015). "Usage-based construction grammar," in *Handbook of Cognitive Linguistics*, eds E. Dabrowska and D. Divjak (Berlin: Mouton de Gruyter)
- Elman, J. L. (1990). Finding structure in time. *Cogn. Sci.* 14, 179–211. doi: 10.1207/s15516709cog1402_1
- Endresen, A. (2015). *Non-standard allomorphy in russian prefixes: corpus, experimental, and statistical exploration* (Doctoral dissertation), UIT The Arctic University of Norway, Tromsø, Norway
- Engelmann, F., Granlund, S., Kolak, J., Szreder, M., Ambridge, B., Pine, J., et al. (2019). How the input shapes the acquisition of verb morphology: elicited production and computational modelling in two highly inflected languages. *Cogn. Psychol.* 110, 30–69. doi: 10.1016/j.cogpsych.2019.02.001
- Ford, A., Singh, R., and Martohardjono, G. (1997). *Pace Panini: Towards a Word-Based Theory of Morphology*. New York, NY: Peter Lang.
- Goldberg, A. (1995). *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago, IL: Chicago University Press.
- Goldberg, A. E. (2002). Surface generalizations: an alternative to alternations. *Cogn. Linguist.* 13, 327–356. doi: 10.1515/cogl.2002.022
- Goldberg, A. E. (2006). *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford University Press.
- Grossberg, S. (1987). Competitive learning: from interactive activation to adaptive resonance. *Cogn. Sci.* 11, 23–63. doi: 10.1111/j.1551-6708.1987.tb00862.x
- Grossberg, S. (2013). Adaptive resonance theory: how a brain learns to consciously attend, learn, and recognize a changing world. *Neural Netw.* 37, 1–47. doi: 10.1016/j.neunet.2012.09.017
- Guenther, F. H., and Perkell, J. S. (2004). "A neural model of speech production and its application to studies of the role of auditory feedback in speech," in *Speech Motor Control in Normal and Disordered Speech*, eds B. Maassen, R. Kent, H. Peters, P. van Lieshout, and W. Hulstijn (Oxford: Oxford University Press), 29–49.
- Harmon, Z. (2019). *Accessibility, language production, and language change* (Doctoral dissertation), University of Oregon, Eugene, OR, United States.
- Harmon, Z., and Kapatsinski, V. (2017). Putting old tools to novel uses: the role of form accessibility in semantic extension. *Cogn. Psychol.* 98, 22–44. doi: 10.1016/j.cogpsych.2017.08.002
- Honey, R. C., Dwyer, D. M., and Illiescu, A. F. (2020). HeiDI: A model for Pavlovian learning and performance with reciprocal associations. *Psychol. Rev.* 127, 829–852. doi: 10.1037/rev0000196
- Idemaru, K., and Holt, L. L. (2011). Word recognition reflects dimension-based statistical learning. *J. Exp. Psychol.* 37, 1939–1956. doi: 10.1037/a0025641
- Inkelas, S. (2015). *The Interplay of Morphology and Phonology*. Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780199280476.001.0001
- Jackendoff, R., and Audring, J. (2016). Morphological schemas: theoretical and psycholinguistic issues. *Ment. Lex.* 11, 467–493. doi: 10.1075/ml.11.3.06jac
- Kahana, M. J. (2002). Associative symmetry and memory theory. *Mem. Cogn.* 30, 823–840. doi: 10.3758/BF03195769
- Kapatsinski, V. (2010a). Velar palatalization in Russian and artificial grammar: constraints on models of morphophonology. *Lab. Phonol.* 1, 361–393. doi: 10.1515/labphon.2010.019
- Kapatsinski, V. (2010b). What is it I am writing? Lexical frequency effects in spelling Russian prefixes: uncertainty and competition in an apparently regular system. *Corpus Linguist. Linguist. Theory* 6, 157–215. doi: 10.1515/cllt.2010.007
- Kapatsinski, V. (2012). "What statistics do learners track? Rules, constraints and schemas in (artificial) grammar learning," in *Frequency Effects in Language Learning and Processing*, eds S. Th. Gries and D. Divjak (Berlin: Mouton de Gruyter), 53–82.
- Kapatsinski, V. (2013). Conspiring to mean: experimental and computational evidence for a usage-based harmonic approach to morphophonology. *Language* 89, 110–148. doi: 10.1353/lan.2013.0003
- Kapatsinski, V. (2014). What is grammar like? A usage-based constructionist perspective. *Linguist. Issues Lang. Technol.* 11, 1–41

- Kapatsinski, V. (2017a). "Learning a subtractive morphological system: statistics and representations," in *Proceedings of the Boston University Conference on Language Development*, Vol. 41, eds M. LaMendola and J. Scott (Somerville, MA: Cascadilla Press), 357–372.
- Kapatsinski, V. (2017b). "Learning morphological constructions," in *The Construction of Words: Advances in Construction Morphology*, ed G. Booij (Cham: Springer), 547–582. doi: 10.1007/978-3-319-74394-3_19
- Kapatsinski, V. (2018). *Changing Minds Changing Tools: From Learning Theory to Language Acquisition to Language Change*. Cambridge, MA: MIT Press. doi: 10.7551/mitpress/11400.001.0001
- Kapatsinski, V., and Harmon, Z. (2017). "A Hebbian account of entrenchment and (over)-extension in language learning," in *Proceedings of the Annual Conference of the Cognitive Science Society*, Vol. 39, eds G. Gunzelmann, A. Howes, T. Tenbrink, and E. J. Davelaar (Austin, TX: Cognitive Science Society), 2366–2371.
- Koranda, M., Zettersten, M., and MacDonald, M. C. (2018). "Word frequency can affect what you choose to say," in *Proceedings of the Annual Conference of the Cognitive Science Society*, Vol. 40, eds T. T. Rogers, M. Rau, X. Zhu, and C. W. Kalish (Austin, TX: Cognitive Science Society), 629–634.
- Langacker, R. (1987). *Foundations of Cognitive Grammar. Vol. I: Theoretical Prerequisites*. Stanford, CA: Stanford University Press.
- Lieberman, P. (1963). Some effects of semantic and grammatical context on the production and perception of speech. *Lang. Speech* 6, 172–187. doi: 10.1177/002383096300600306
- Matzel, L. D., Held, F. P., and Miller, R. R. (1988). Information and expression of simultaneous and backward associations: implications for contiguity theory. *Learn. Motiv.* 19, 317–344. doi: 10.1016/0023-9690(88)90044-6
- McClelland, J. L., and Elman, J. L. (1986). The TRACE model of speech perception. *Cogn. Psychol.* 18, 1–86. doi: 10.1016/0010-0285(86)90015-0
- Neset, T. (2008). *Abstract Phonology in a Concrete Model: Cognitive Linguistics and the Morphology-Phonology Interface*. Berlin: Walter de Gruyter. doi: 10.1515/9783110208368
- Nixon, J. S. (2020). Of mice and men: speech sound acquisition as discriminative learning from prediction error, not just statistical tracking. *Cognition* 197:104081. doi: 10.1016/j.cognition.2019.104081
- Okada, K., Matchin, W., and Hickok, G. (2018). Neural evidence for predictive coding in auditory cortex during speech production. *Psychon. Bull. Rev.* 25, 423–430. doi: 10.3758/s13423-017-1284-x
- Onnis, L., Waterfall, H. R., and Edelman, S. (2008). Learn locally, act globally: learning language from variation set cues. *Cognition* 109, 423–430. doi: 10.1016/j.cognition.2008.10.004
- O'Reilly, R. C., and Munakata, Y. (2000). *Computational Explorations in Cognitive Neuroscience: Understanding the Mind by Simulating the Brain*. Cambridge, MA: MIT Press. doi: 10.7551/mitpress/2014.001.0001
- Osgood, C. E. (1963). On understanding and creating sentences. *Am. Psychol.* 18, 735–751. doi: 10.1037/h0047800
- Pereira, A. F., Smith, L. B., and Yu, C. (2014). A bottom-up view of toddler word learning. *Psychon. Bull. Rev.* 21, 178–185. doi: 10.3758/s13423-013-0466-4
- Pinker, S., and Prince, A. (1988). On language and connectionism: analysis of a parallel distributed processing model of language acquisition. *Cognition* 28, 73–193. doi: 10.1016/0010-0277(88)90032-7
- Ramscar, M., Hendrix, P., Shaoul, C., Milin, P., and Baayen, H. (2014). The myth of cognitive decline: non-linear dynamics of lifelong learning. *Top. Cogn. Sci.* 6, 5–42. doi: 10.1111/tops.12078
- Ramscar, M., Yarlett, D., Dye, M., Denny, K., and Thorpe, K. (2010). The effects of feature-label-order and their implications for symbolic learning. *Cogn. Sci.* 34, 909–957. doi: 10.1111/j.1551-6709.2009.01092.x
- Rescorla, R. A., and Wagner, A. R. (1972). "A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement," in *Classical Conditioning II: Current Research and Theory*, eds A. H. Black and W. F. Prokasy (New York, NY: Appleton Century Crofts) 64–99.
- Roelofs, A. (1999). Phonological segments and features as planning units in speech production. *Lang. Cogn. Process* 14, 173–200. doi: 10.1080/016909699386338
- Roembke, T. C., Wasserman, E. A., and McMurray, B. (2016). Learning in rich networks involves both positive and negative associations. *J. Exp. Psychol.* 145, 1062–1074. doi: 10.1037/xge0000187
- Shih, S. S. (2017). "Phonological influences in syntactic alternations," in *The Morphosyntax-Phonology Connection: Locality and Directionality at the Interface*, eds V. Gribanova and S. S. Shih (Oxford: Oxford University Press). doi: 10.1093/acprof:oso/9780190210304.003.0009
- Smolek, A. E. (2019). *Teaching papa to cha-cha: how change magnitude, temporal contiguity, and task affect alternation learning* (Doctoral dissertation), University of Oregon, Eugene, OR, United States.
- Taylor, J. R. (2012). *The Mental Corpus: How Language is Represented in the Mind*. Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780199290802.001.0001
- Tiersma, P. M. (1982). Local and general markedness. *Language* 58, 832–849. doi: 10.2307/413959
- Wang, H. S., and Derwing, B. L. (1994). "Some vowel schemas in three English morphological classes: experimental evidence," in *In Honor of Professor William S.-Y. Wang: Interdisciplinary Studies on Language and Language Change* eds M. Y. Chen and O. J. L. Tseng (Taipei: Pyramid Press), 561–575.
- Xu, F., and Tenenbaum, J. B. (2007). Word learning as Bayesian inference. *Psychol. Rev.* 114, 245–272. doi: 10.1037/0033-295X.114.2.245
- Zaliznjak, A. A. (1974). *Obratnyj slovar' russkogo jazyka: okolo 125000 slov*. Moscow: Sovetskaja Enciklopedija.

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Kapatsinski. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.