Language and Cognition **8** (2016), 369–390. doi:10.1017/langcog.2016.17
© UK Cognitive Linguistics Association, 2016

Partial productivity of linguistic constructions: Dynamic categorization and statistical preemption*

ADELE E. GOLDBERG

Psychology Department, Princeton University

(*Received* 01 *July* 2015 – *Revised* 01 *November* 2015 – *Accepted* 07 *March* 2016 – *First published online* 14 *July* 2016)

ABSTRACT

Grammatical constructions are typically partially but not fully productive, which leads to a conundrum for the learner. When can a construction be extended for use with new words and when can it not? The solution suggested here relies on two complementary processes. The first is DYNAMIC CATEGORIZATION: as learners record the statistics of their language, they implicitly categorize the input on the basis of form and function. On the basis of this categorization process, general semantic and phonological constraints on productivity emerge, and productivity is to a large extent determined by the degree to which the category is well attested by similar exemplars. Occasionally, a semantically sensical and phonologically well-formed instance of a well-attested construction is simply not fully acceptable. It is suggested that a process of STATISTICAL PREEMPTION is at work in these cases: learners avoid using a construction if an alternative formulation has been systematically witnessed instead. The mechanism proposed for statistical preemption is competition-driven learning: when two competitors are activated but one reliably wins, the loser becomes less accessible over time. In this way, the paradox of partial productivity can be resolved.

KEYWORDS: productivity, constructions, induction.

1. Introduction

A learner's goal is to understand intended messages given the particular forms that are witnessed for the sake of comprehension, and to choose particular forms, given the intended information she wishes to convey for the sake of production. Therefore it is clear that speakers must learn the

[*] I am grateful to Andrea Tyler and an anonymous reviewer for helpful comments on an earlier draft, and to Ben Ambridge, Jeremy Boyd, Clarice Robenalt, and Laura Suttle for much discussion on the topics presented here. e-mail: adele@princeton.edu

https://doi.org/10.1017/langcog.2016.17 Published online by Cambridge University Press

ways in which forms and functions are paired in the languages they speak. These learned pairings of form and function are referred to here as CONSTRUCTIONS. Constructions are understood to vary in their degree of complexity and abstraction, and to form an inter-related dynamic network of linguistic knowledge. A few English constructions are provided in Table 1, along with exemplars of each, attested in the Corpus of Contemporary American English (COCA: Davies, 2008a).[1]

The ability to cluster – DYNAMICALLY CATEGORIZE – witnessed exemplars into distributions of types is clearly ubiquitous in humans and throughout the animal kingdom. For example, the next door we encounter may differ from previous doors in being larger or smaller, wooden or windowed, and may require pushing, pulling, or sliding to open. And yet we have no trouble recognizing a new door as a door; nor, fortunately, do we normally have trouble distinguishing doors from windows. We categorize linguistic elements as well (e.g., Kuhl, 2000; Lakoff, 1987; Langacker, 1987; Taylor, 2003). As discussed below, each construction forms a category, and this allows us to apply our linguistic knowledge to new situations and experiences. That is, constructions are PRODUCTIVE to varying degrees. A few examples of productive uses of familiar constructions (again labeled on the right) are provided in Table 2.

At the same time, the same constructions exemplified in Tables 1 and 2 resist being used productively with certain other words, even when the intended meaning is perfectly clear and the examples do not violate system-wide semantic, syntactic, or phonological generalizations. Examples that illustrate the lack of full productivity are provided in Table 3, along with related acceptable examples in parentheses.

Thus, constructions are typically partially productive in that they can be extended for use with some words (Table 2), but they are not necessarily completely productive, even when no general semantic, phonological, or syntactic constraints are violated (Table 3). The present paper investigates the long-standing paradox that this partial productivity presents: How do learners know when and how far a given construction's productivity extends?

A good deal of work has demonstrated that the solution is non-trivial. Learners do not reliably receive overt corrections for ill-formed utterances, because people are much more interested in the content of a speaker's contribution than its form (Baker, 1979; Bowerman, 1988, 1996; Braine, 1971; Brown & Hanlon, 1970; Marcus, 1993; Pinker, 1989). That the words used 'fit' the constraints on the construction is required, as explained in Section 2 (see also Ambridge, Pine, Rowland, Jones, & Clark, 2009; Coppock, 2008;

[1] Unless otherwise specified, all examples in quotes are from COCA, a free, parsed, roughly 450 million word corpus of spoken and written texts made available on-line by Mark Davies: <http://view.byu.edu/> (Davies, 2008a).

https://doi.org/10.1017/langcog.2016.17 Published online by Cambridge University Press

TABLE 1. *Four English* CONSTRUCTIONS *(learned pairings of form and function) and exemplars of each from COCA*

Exemplar	Construction
(1) "Janson gave him a business card."	Double-object construction
(2) "We broke the shredder."	Transitive causative construction
(3) "She had hoped to look enticing."	To-infinitival complement construction
(4) "the tall figure"	Attributive modification construction

TABLE 2. *Novel linguistic exemplars that demonstrate the* PRODUCTIVITY *of various constructions*

Exemplar	Construction
(5) "Hey man, bust me some fries." (Google)	Double-object construction
(6) "We were wardrobing her for her nightly chat show" (Google)	Transitive causative construction
(7) "I IM'ed him to go ahead and PayPal the funds" (Google)	To-infinitival complement construction
(8) "what a bodacious thing to say" (Google)	Attributive modification construction

TABLE 3. *Novel formulations that are judged odd by native speakers*

Exemplar	Construction
(9) ??"She explained him the story." (cf. "She told/guaranteed him the story.")	Double-object construction
(10) ??"He vanished the rabbit." (cf. "He banished the rabbit.")	Transitive causative construction
(11) ??"She considered to go." (cf. "She hoped/planned to go.")	To-infinitival complement construction
(12) ??"the asleep boy" (cf. "the astute/sleeping boy")	Attributive modification construction

Goldberg, 1995; Gropen, Pinker, Hollander, & Goldberg, 1991; Gropen, Pinker, Hollander, Goldberg, & Wilson, 1989; Pinker, 1989), but it is not sufficient to insure acceptability, as illustrated in the examples in Table 3. Positing underlying or invisible features does not address the learning issue, since doing so would beg the question of how it is that learners know to assign the relevant diacritics to some lexical items and not others (Ambridge, Pine, & Lieven, 2015; Goldberg, 2011b; Pinker, 1989, section 5.2).

It is tempting to believe that speakers only use familiar words in the ways in which they have been witnessed, i.e., that speakers are wholly CONSERVATIVE (Baker, 1979; Braine & Brooks, 1995). In line with this idea, it has been predicted that the more often a word is witnessed in one construction, the more difficult

https://doi.org/10.1017/langcog.2016.17 Published online by Cambridge University Press

it is to extend it for use in a different construction (Ambridge et al., 2009; Stefanowitsch, 2008). In fact, children are relatively more willing to overgeneralize infrequent verbs (e.g., to use *vanish* transitively) than to overgeneralize frequent verbs (e.g., to use *disappear* transitively) (Ambridge, Pine, & Rowland, 2012; Theakston, 2004). The suggestion has been that this is due to the fact that *disappear* has been heard in the simple intransitive construction much more often than *vanish*, and that it is more difficult to creatively causativize because it is more entrenched intransitively. We revisit this finding in Section 3.

This proposal, which is referred to here as CONSERVATIVISM VIA ENTRENCHMENT, faces a problem, because if learners only use predicates in ways in which they have already been witnessed, and if predicates more strongly resist novel uses for higher-frequency verbs, then the following attested examples ought to be quite ill-formed:

(13) [she] prayed her way through the incomprehension of her atheist friends
(14) The python coughed her back out (<www.rabbit.org/journal/3-7/snake-bite.html>)
(15) Aladar [a dinosaur] swam his friends to the mainland. (Disney, *Aladar*)
(16) He's right here at my feet, snoring his head off.

Each of the verbs in (13)–(16) (*pray*, *cough*, *swim*, *snore)* is very frequent ('entrenched') in the INtransitive construction, and only exceedingly rarely, if ever, witnessed in the various transitive constructions in (13)–(16).[2] And yet, although Robenalt and Goldberg (2015) find that such novel sentences are in fact judged to be less acceptable than sentences in which the same verbs are used intransitively, they are not as ill-formed as the types of novel examples in Table 3. Moreover, speakers readily extend verbs in new ways that have not been witnessed when the intended message is conveyed better by a different construction (Perek & Goldberg, 2015). Thus, the solution to the issue of partial productivity is not merely a matter of learners being conservative via entrenchment.

In the following sections, it is argued the solution follows from the fact that attested exemplars cluster together to form constructional categories, and that constructions can compete with one another in particular contexts. A concrete example may be helpful. If we learn that many varieties of leafy green vegetables are called *lettuce*, we are likely to label a new, only subtly different, leafy green vegetable as *lettuce* as well. That is, if we know that a category is attested by a variety of exemplars, and a new exemplar is sufficiently similar to attested instances, we are very likely to assign it to the same category. At the same time,

[2] For example, *pray* occurs 7,929 times in COCA, but only five of those are examples of the *way* construction as in (13).

https://doi.org/10.1017/langcog.2016.17 Published online by Cambridge University Press

if we hear a different label, say *kale,* consistently assigned to a new type of leafy green vegetable in contexts in which we might have expected to hear *lettuce*, then we will learn that kale is not lettuce (see also Bowerman & Choi, 2003).

Briefly, the analogy to syntactic productivity outlined in more detail below is as follows. A potential productive use of an existing construction (a new COINAGE) is acceptable to the extent that the extended category that includes previously attested examples and the potential coinage is well attested (i.e., is dense or well-COVERED). The idea that speakers generalize over attested exemplars suggests that semantic, pragmatic, and phonological constraints emerge, as exemplars that share the same surface form are categorized. For example, exemplars of the English double-object formal pattern construction will almost all share an implication of transfer from one entity to another, and they will almost always involve a more topical recipient argument and a more focal theme argument. As these exemplars are categorized as instances of the same construction, the well-known semantic and information structure constraints of the double-object construction will emerge.

At the same time, as we saw in Table 3, there are certain formulations that are avoided by native speakers even though they seem to fit within these types of emergent constraints. It is proposed that a new coinage will be inhibited to the extent that there already exists a readily available alternative formulation that serves the requisite function; in this case, the alternative will STATISTICALLY PREEMPT the coinage. To return to our lettuce example, the category of *lettuce* is well attested by a variety of exemplars, all of which are leafy green vegetables. But, since a particular type of leafy green is consistently labeled *kale* in contexts where one might have expected to hear *lettuce*, people learn that that type of leafy green is kale and not lettuce. In Sections 2 and 3, these two aspects of the proposal, coverage – which encourages productivity while capturing emergent semantic and phonological generalizations – and statistical preemption – which constrains productivity and accounts for the learning of seemingly arbitrary exceptions – are discussed in turn.

2. The range of generalization is determined by coverage

Work by Suttle and Goldberg (2011) and Perek (2016) has argued that the critical factor in determining when a construction is productive is COVERAGE, an idea borrowed from the non-linguistic categorization literature (Goldberg, 2006, p. 98; Osherson, Smith, Wilkie, Lopez, & Shafir, 1990). Coverage relates type frequency, variability, and similarity of the coinage to attested tokens: all factors that have been independently found to be relevant. The idea is

https://doi.org/10.1017/langcog.2016.17 Published online by Cambridge University Press

depicted in Figure 1. A new coinage is acceptable to the extent that the semantic (pragmatic, and/or phonological) space is well covered by the smallest convex category that encompasses both the coinage and attested instances that share the same formal pattern: the category is represented by the larger oval. Exemplars with shared form are represented in a high degree similarity space, projected here onto two dimensions for expository purposes. The degree of coverage corresponds to the degree to which the attested instances fill or 'cover' the entire category.

In a series of experiments performed using Amazon's Mechanical Turk, Suttle and Goldberg (2011) found that type frequency, variability of attested instances, and similarity of a target utterance to attested instances interact in ways that are predicted by the notion of coverage. The design of the experiment was as follows. We provided one to six attested utterances of a fictitious language, Zargotian, and then asked participants to judge how likely it was that a final utterance would also be acceptable in Zargotian. As example stimulus trial is given below:

(17) Assume you can say these sentences.
Scrape-*nu* the *vip* the *hap*.
Load-*nu* the *yib* the *vork*.
Flip-*nu* the *loof* the *rolm*.
How likely is it, on a scale of 1–100, that you can also say:
Rumple-*nu* the *pheb* the *jirm*.

We systematically varied (i) whether participants were given one, three, or six distinct attested exemplars (type frequency), (ii) the diversity of verb classes the exemplars were chosen from (variability), and (iii) the degree of similarity between the target utterance and its closest attested neighbor, as determined by Latent Semantic Analysis (Landauer, 2006). Ten verb classes were varied across participants and items and included verbs of breaking, loading, bending, cooking, cutting, acquiring, throwing, hitting, holding, and cognition.

The findings confirmed that when coverage is relatively high, a coinage is judged to be more acceptable. For example, in the situation depicted in Figure 2, in which three attested examples come from different verb classes and the potential coinage comes from yet a different class, participants judged the potential coinage to be less acceptable than if type frequency was increased and all else was held constant (as depicted in Figure 3).

If a new coinage is sufficiently semantically dissimilar so that coverage is again low, the coinage is judged less acceptable, even if the type frequency and variability of attested instances is relatively high (Suttle & Goldberg, 2011, experiment 3). This situation is depicted in Figure 4 (see also Barðdal, 2008; Bybee & Eddington, 2006; Croft & Cruse, 2004; Kalyan, 2012; Langacker, 1987; Wonnacott, Boyd, Thompson, & Goldberg, 2012; Zeschel & Bildhauer, 2009).

https://doi.org/10.1017/langcog.2016.17 Published online by Cambridge University Press

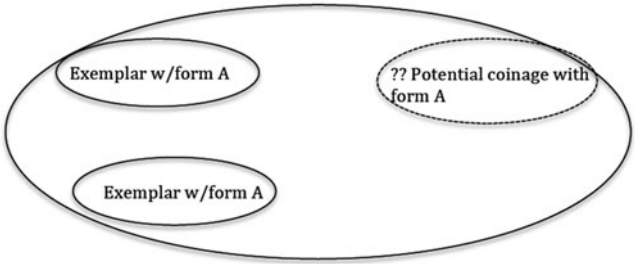

Fig. 1. The smallest convex category in similarity space that includes both attested examples and a potential coinage. The extent to which the instances COVER the category correlates with how acceptable the coinage is judged to be.

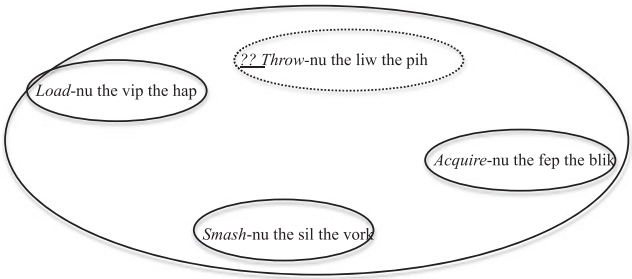

Fig. 2. Sample stimuli involving relatively low coverage from Suttle and Goldberg (2011, experiment 1), represented pictorially.

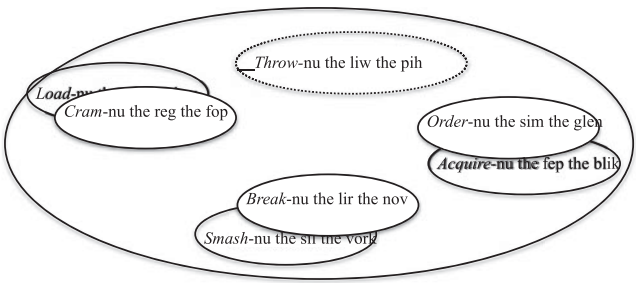

Fig. 3. Sample stimuli involving higher coverage than that depicted in Figure 2 due to higher type frequency, from Suttle and Goldberg (2011, experiment 2) represented pictorially.

The role of type frequency interacts with semantic similarity in the following way. If the potential coinage is semantically similar to a cluster of examples with high type frequency and high semantic similarity, then the coinage is likely to be judged quite acceptable. However, acceptability decreases as the semantic similarity of the potential coinage to the cluster decreases. Thus, a lack of semantic variability of attested tokens inhibits

https://doi.org/10.1017/langcog.2016.17 Published online by Cambridge University Press

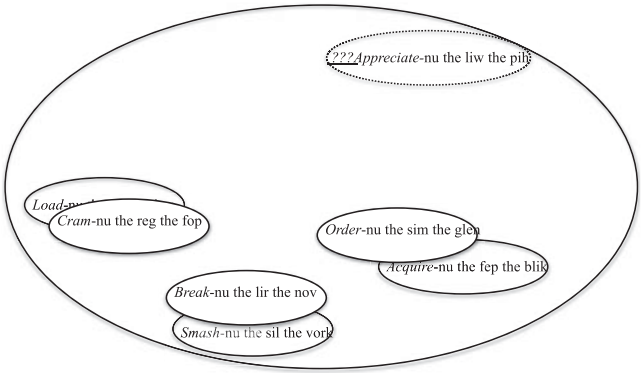

Fig. 4. Type frequency and variability is the same as is represented in Figure 3, and yet coverage is reduced because the potential coinage is less similar to the attested types.

generalization if the potential coinage is not part of the same cluster of related tokens, as depicted in Figure 5. This type of relationship between type frequency and variability has also been reported previously (Barðdal, 2008; Bowerman & Choi, 2001; Bybee, 1985, 1995; Clausner & Croft, 1997; Goldberg, 1995; Janda, 1990; Tomasello, 2003; Xu & Tenenbaum, 2007).

Thus the notion of COVERAGE is a way of combining the well-supported and independently recognized factors of type frequency, variability, and similarity of a potential coinage to attested exemplars. Support for the notion of coverage comes from Perek (2016), who investigates the nature of productivity over time by examining the 'V the hell out of NP' construction exemplified in (18).

(18) Santas that would scare the hell out of Jesus. (Google)

He examines the semantic distribution of verbs used in the construction in each of four 20-year time periods between 1930 and 2009, using distributional semantics and multidimensional scaling on the attested verbs found in COHA (Davies, 2008b). Perek's results demonstrate that the degree of density of a semantic cluster during one period strongly correlates with how many new verbs are added to the cluster in the following two decade time period. That is, clusters with higher density tend to attract near neighbors to their cluster, just as the notion of coverage predicts.[3]

[3] The role of token frequency and its interaction with type frequency requires much more study; I leave this issue aside for now (but see, e.g., Boyd & Goldberg, 2009; Bybee, 1985, 1995, 2010; Casenhiser & Goldberg, 2005; Desagulier, 2015; Ellis & Ferreira-Junior, 2009; Goldberg, Casenhiser, & Sethuraman, 2004; Hilpert, 2013; Madlener, 2015; McDonough & Nekrasova-Becker, 2014; Wonnacott et al., 2012).

https://doi.org/10.1017/langcog.2016.17 Published online by Cambridge University Press

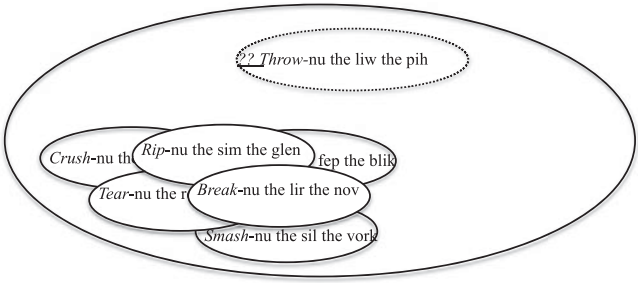

Fig. 5. High type frequency does not increase coverage if the potential coinage falls outside the similarity space defined by attested tokens.

Categorization, as captured by the notion of coverage, thus allows for the fact that language is often productive within a circumscribed semantic, pragmatic, and phonological space. That is, coverage captures the idea that new uses of verbs must fit, or be able to accommodate, the semantic, pragmatic, and phonological constraints of the constructions they appear in (Ambridge et al., 2009; Coppock, 2008; Goldberg, 1995; Gropen et al., 1989; Gropen, et al., 1991; Pinker, 1989). Since speakers implicitly categorize instances of each construction, and thereby form generalizations about semantic, pragmatic, and phonological constraints, new expressions are judged to be well-formed to the extent that they satisfy the general constraints of the constructions involved.

At the same time, coverage is not sufficient in itself to account for the actual distribution of acceptable and non-acceptable exemplars. Recalling the examples in Table 3, it is clear that certain exemplars are ill-formed, even though they satisfy the general constraints on the constructions in question. That is, attested instances of the constructions involved appear to cover the similarity space that should include the examples in Table 3, and yet these examples nonetheless sound odd to native speakers.

3. Statistical preemption: competition-dependent learning

How is it that children learn to avoid the unacceptable examples in Table 3? This question has bedeviled researchers for decades (Ambridge, Pine, & Rowland, 2012; Ambridge, Pine, Rowland, & Young, 2008; Ambridge et al., 2009; Baker, 1979; Braine, 1971; Bowerman, 1988; Goldberg, 1995, 2006, 2011a; Pinker, 1989). In this section, it is argued that a process of STATISTICAL PREEMPTION plays a key role (Clark, 1987; Foraker, Regier, Khetarpal, Perfors, & Tenenbaum, 2007; Goldberg 1993, 1995, 2006, 2011a; Marcotte, 2005). Statistical preemption is a particular type of indirect negative evidence

https://doi.org/10.1017/langcog.2016.17 Published online by Cambridge University Press

that results from repeatedly hearing a formulation, B, in a context where one might have expected to hear a semantically and pragmatically related alternative formulation, A. Given this type of input, speakers recognize that B is the appropriate formulation in such a context, and implicitly learn that A is not appropriate.

Morphological preemption (or 'blocking') has long been familiar from morphology: *went* preempts *goed*, and *feet* preempts *foots* (Aronoff, 1976; Kiparsky, 1993; Rainer, 1988). That is, children learn to produce *feet* instead of *foots* because they systematically hear *feet* every time the 'plural of *foot*' is expressed. At the same time, in the case of phrasal constructions, the role of statistical preemption requires discussion, since, unlike *feet* and the potential *foots*, distinct phrasal constructions are virtually never semantically and pragmatically identical (Bolinger, 1977; Clark, 1987; Goldberg, 1995). Since two constructions that are semantically related often happily co-occur with the same verb, some have argued that statistical preemption cannot be effective (Bowerman, 1996; Pinker, 1989). Certainly, knowledge that the *to*-dative paraphrase is licensed for *explain* should not immediately preempt the use of the double-object construction, since a large number of verbs freely appear in both constructions (e.g., *tell*).

But the fact that each construction has a distinct function can actually work in favor of statistical preemption. Consider the *to*-dative and double-object constructions. They have overlapping, but distinct, semantic and information structure properties in that many corpus and production studies have demonstrated that the double-object construction is preferred over the *to*-dative if the recipient argument is pronominal and the transferred entity is a lexical noun phrase (Arnold, Eisenband, Brown-Schmidt, & Trueswell, 2000; Bresnan, Cueni, Nikitina, & Baayen, 2007; Collins, 1995; Dryer, 1986; Erteschik-Shir, 1979; Givón, 1979, 1984; Goldberg, 1995, 2006; Green, 1974; Oehrle, 1975; Thompson, 1990, 1995; Wasow, 2002). For instance, examples like (19) are vastly more common than those like (20).

(19) She gave me the ball.
(20) She gave the ball to me.

The difference between the double-object and *to*-dative constructions is subject to some dialect differences and gradability, yet it is possible to predict with high probability which construction will be preferred in a given context, for a given dialect (Bresnan & Ford, 2010; Bresnan & Hay, 2008). Therefore learners will witness situations in which the double-object construction is expected for a given verb, because the relevant information structure suits the double-object construction at least as well as the *to*-dative. If, in these situations, the *to*-dative is systematically witnessed instead, the learner can

https://doi.org/10.1017/langcog.2016.17 Published online by Cambridge University Press

infer that the double-object construction is not after all appropriate (Goldberg, 1995, 2006, 2011a). As Goldberg (2006) emphasizes, the process is necessarily statistical, because a single use of the *to*-dative could be due to an unrecognized factor that actually encourages the *to*-dative, or even to an error by the speaker. But if the *to*-dative is consistently heard in such contexts, statistical preemption will lead to an avoidance of the double-object construction in favor of the *to*-dative. More generally, because of the difference in function between two constructions, A and B, there will exist contexts in which A is at least as appropriate as B for a particular verb. If B is consistently witnessed instead, people can learn that A is not possible for that verb.

Statistical preemption of phrasal forms has been investigated experimentally in only a few studies. Brooks and colleagues have found that novel intransitive verbs that have been witnessed in the preemptive periphrastic causative construction are much less likely to be used in the simple transitive than those that have not (Brooks & Tomasello, 1999; Brooks, & Zizak 2002). For example, if a child hears both *The cow is chamming* and *Ernie's making the cow cham*, they are less likely to respond to "What did Elmo do to the cow?" with *Ernie chammed the cow* (the causative), than they are if only the intransitive construction had been witnessed (Brooks & Tomasello, 1999). It seems that hearing the novel verb used in the periphrastic causative construction provides a readily available alternative to the causative construction, statistically preempting the use of the latter (cf. also Tomasello, 2003).

Another case of an unpredictable restriction involves certain adjectives such as *afraid* which resist prenominal attributive position (21a), despite the fact that near synonyms and phonologically analogous adjectives readily appear in this position (21b):

(21) a. ??the afraid boy
b. the scared/aloof boy

These *a*-adjectives begin with an unstressed schwa and can be morphologically segmented into *a*- plus a semantically related stem (e.g., *a-live*, *a-sleep*). The distribution is motivated by the fact that the majority of *a*-adjectives historically were prepositional phrases and, as prepositional phrases, they could not be expected to appear prenominally. Like typical adjectives, *a*-adjectives are inseparable phonological units, modify nouns, can be conjoined with uncontroversial adjectives (22) and can appear after the verb *seem* (23):

(22) The man was quiet and afraid.
(23) The man seemed afraid/asleep.

Thus, since speakers are generally unaware of the historical facts, the question arises as to how the restriction can be learned.

https://doi.org/10.1017/langcog.2016.17 Published online by Cambridge University Press

Boyd and Goldberg (2011) examined adult naturalistic productions of such adjectives in three experiments, all of which required participants to describe scenes in which one of two animals with different adjective labels moved to a star. The experiments all included four classes of adjectives: real *a*-adjectives; nearly synonymous real non-*a*-adjectives; nonsense *a*-adjectives; and nonsense non-a-adjectives. The task resulted in either a relative clause or prenominal (attributive) use of the target adjective (e.g., (24) or (25)).

(24) Prenominal:
The sleepy/??asleep/?adax fox.
(judgments based on data from Experiment 1 of Boyd & Goldberg, 2011)
(25) Relative clause:
The fox that's sleepy/asleep/adax.

The first experiment established that real *a*-adjectives (e.g., *asleep*) strongly disprefer prenominal use, relative to non-*a* adjectives (e.g., *sleepy*). In addition, NOVEL *a*-adjectives (e.g., *adax*) disprefer prenominal use relative to non-*a* adjectives (e.g., *chammy*) to a significant extent as well. This indicates that participants tentatively assimilate never-before-seen *a*-adjectives to the category of familiar *a*-adjectives. The real *a*-adjectives were much less likely to occur prenominally than the novel *a*-adjectives were, but it suggests that speakers can TENTATIVELY generalize a restriction to unwitnessed but similar exemplars.

A second experiment investigated the role of statistical preemption. It was found that in fact witnessing two of the four novel *a*-adjectives used in a preemptive relative clause context just three times each dramatically decreased prenominal uses so that all four novel *a*-adjectives behaved indistinguishably from familiar *a*-adjectives in avoiding prenominal uses. Non-*a*-adjectives were unaffected. This result is striking because it not only demonstrates the effectiveness of preemption, but it also demonstrates that speakers are able to generalize evidence gleaned from statistical preemption to other members of the same category.

A final experiment showed that learners rationally disregard PSEUDO-preemptive input. Speakers did not display an increased avoidance of prenominal uses when exposed to pseudo-preemptive contexts like (26), presumably because they rationally attributed *adax*'s appearance in the relative clause to the complex adjective (cf. (27)), rather than to *adax*.

(26) The hamster, adax and proud of itself, moved to the star.
(27) *The proud of itself hamster moved to the star.

Productions in the last experiment patterned like those in the first experiment where no preemptive context was provided. Fillers were used to obscure the goal of the experiment and to guard against the effects being a simple result

https://doi.org/10.1017/langcog.2016.17 Published online by Cambridge University Press

of structural priming. Debriefing confirmed that speakers were unaware of the manipulations (see Goldberg & Boyd, 2015, Yang, 2015, for further discussion).

Collectively, these experiments go some way toward establishing how speakers are able to learn arbitrary distributional restrictions in their language – i.e., how they learn what not to say. Learners categorize their input, tentatively generalizing restrictions to new members of a perceived category. Familiar formulations statistically preempt other formulations when the former are repeatedly witnessed instead of a hypothesized formulation. Providing evidence that speakers categorize restrictions, the second experiment demonstrated that speakers extended the information gained from preemptive contexts to other instances of the same category. At the same time, speakers use statistical preemption wisely: they are impressively adept at ignoring alternative formulations when those formulations can be attributed to some irrelevant factor.

The preemptive process, unlike the notion of conservatism via entrenchment, predicts that expressions like (13)–(16) would NOT be preempted by the overwhelmingly more frequent uses of *pray*, *cough*, *swim*, and *snore* intransitively because the expressions in (13)–(16) are not in competition with the intransitive uses. For example, the meanings of causing a change of state (28) and an involuntary intransitive action (29) would not be used in the same contexts:

(28) And he sneezed the house in! (Joseph Robinette, *The trial of the big bad wolf*)
(29) She sneezed.

The intriguing finding that high-frequency intransitive verbs (e.g., *disappear NP*) are less acceptable when used causatively than low-frequency intransitive verbs (e.g., *vanish NP*) is consistent with the idea that it is preemption that prevents overgeneralization, rather than the frequency of the verb per se. Note that the periphrastic causative of high-frequency verbs is more frequent than that of low-frequency verbs. In fact, a corpus search of the Corpus of Contemporary American English confirms that (30) is more frequent than (31), by a factor of ten.

(30) [NP] made [NP] disappear.
(statistically preempts [NP disappeared NP])
(31) [NP] made [NP] vanish.
(statistically preempts [NP vanished NP])

Robenalt and Goldberg (2015) revisit the finding that lower-frequency verbs are more acceptable in novel constructions, relative to their baseline acceptability in familiar types of sentences. If it is preemptive expressions that lead to the novel uses of the verbs being judged unacceptable, rather

https://doi.org/10.1017/langcog.2016.17 Published online by Cambridge University Press

than baseline expressions, we should not find the same frequency effect for those novel expressions that do not have a readily available alternative. To see whether this prediction held, pairs of novel sentences were created, each involving both low- and high-frequency near-synonyms, with novelty confirmed using the COCA corpus (Davies, 2008a). In a separate norming study, the sentence pairs were classified into two groups according to whether there exists a readily available paraphrase. Specifically, if more than half of a group of naive participants suggested THE SAME paraphrase for a given sentence, the sentence was considered to have a competing alternative; if instead no single paraphrase was agreed upon by the majority of participants, the sentence was considered not to have a readily available competing alternative. For example, in response to (32), the majority of respondents suggested the same alternative: *Natalie smacked the mosquito with a newspaper*. On the other hand, in the case of (33), people instead proposed a wide variety of paraphrases, e.g., *The magician was so fascinating the toddlers went into a trance; The magician entertained the toddlers and they became fascinated*, etc.

(32) Natalie smacked a newspaper onto the mosquito.
(33) The magician fascinated the toddlers into a trance.

Thus (32) has a readily available competing alternative and (33) does not.

Findings replicated the stronger dispreference for a novel use with a high-frequency verb relative to its lower-frequency counterpart, but only for those sentences with a competing alternative phrasing. That is, while *smack* is judged worse than *swatted* in the caused motion construction (*Natalie smacked/swatted a newspaper onto the mosquito*), frequency had no effect on novel sentences that had no readily available alternative, such as (33) or (13)–(16). For example, despite the fact that *fascinate* is more frequent than *enthrall*, the sentence *The magician fascinated the toddlers into a trance* was not judged to be less acceptable than *The magician enthralled the toddlers into a trance*. Thus, when there is no consensus about a preferred way to phrase a sentence, verb frequency is not a predictive factor in a sentence's ratings. This result implies that speakers are not simply conservative overall – they are willing to extend familiar words in new ways, but they ARE conservative when a readily available alternative formulation already exists. When it does, the readily available formulation is preferred – and the strength of the preference varies with the frequency of the competing alternative. Thus witnessing exemplars of one construction and NOT exemplars of a competing construction can lead learners to judge the non-occurring form to be unacceptable. This is represented schematically in Figure 6.

If a novel formulation is not in competition with a familiar formulation, additional evidence of the familiar formulation does not weigh against the use of the novel formulation (Figure 7).

https://doi.org/10.1017/langcog.2016.17 Published online by Cambridge University Press

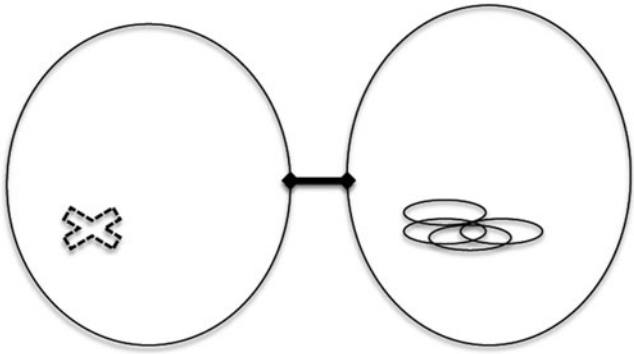

Fig. 6. Two competing constructions (competition indicated by the solid bar linking them). Attested instances on the right serve to statistically preempt the productive use on the left (indicated by the cross).

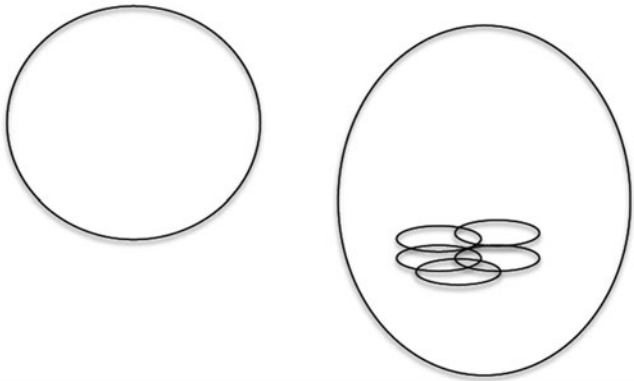

Fig. 7. If there is no competition between two constructions, witnessing instances of one has no bearing on whether a novel instance of the other is judged acceptable.

This is not to say that the degree of familiarity is irrelevant. Robenalt and Goldberg (2015) found that, overall, sentences in which verbs were used in their familiar argument structure pattern were strongly preferred over novel formulations, whether there existed a readily available alternative to the novel sentences or not (see also work by Ambridge and colleagues, e.g., Ambridge et al. 2012).[4] We can thus summarize the results as follows. Speakers prefer to use the types of exemplars they have witnessed in

[4] Robenalt and Goldberg (2016) replicated this result in a separate group of native speakers, and also found that L2 learners only pattern with native speakers at the highest quartile of proficiency. Possible factors that lead to the difference between L1 and lower proficiency L2 speakers are explored in that paper.

https://doi.org/10.1017/langcog.2016.17 Published online by Cambridge University Press

the input, but they are willing to extend constructions productively unless there exists a readily available alternative way of expressing the intended meaning.

3.1. MECHANISM: COMPETITION-DRIVEN LEARNING

There is a great deal of evidence that we often predict what others will say as they speak (e.g., Johnson, Turk-Browne, & Goldberg, 2013; Kutas & Hillyard, 1984; McRae, Spivey-Knowlton, & Tanenhaus, 1998; Pickering and Garrod, 2007, 2013; Stephens, Silbert, & Hasson, 2010; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). When speakers anticipate a particular construction, we can assume that the construction is PARTIALLY ACTIVATED. Intriguingly, it turns out that if one representation is partially activated, but a competing form is accessed instead, the partially activated form is subsequently harder to retrieve. This is true at the level of individual neurons: strong excitatory input leads to long-term synaptic strengthening, but MODERATE excitatory input leads to long-term synaptic weakening (Artola, Brocher, & Singer, 1990).

Behaviorally, too, partial activation of a competing form leads to learned dissociation (Anderson, Green, & McCulloch 2000; Anderson & Spellman, 1995; Kim, Lewis-Peacock, Norman, & Turk-Browne, 2014; Newman & Norman, 2010; Norman, Newman, & Detre, 2007; Storm & Levy, 2012). The effect, often referred to as RETRIEVAL INDUCED FORGETTING, has been demonstrated, for example, in the following type of paradigm. Anderson and Spellman (1995) had a group of subjects learn paired associations, e.g., Fruit–Apple, Fruit–Pear, Fruit–Kiwi, Furniture–Table, Sport–tennis, Furniture–Chair, and so on. Participants were then provided incomplete cues in order to retrieve a subset of these pairs. For instances, one incomplete cue had the form:

(34) Fruit-Pe___.

Note that since 'Pear' is only partially cued in (34), subjects can be expected to partially activate other prototypical associates of Fruit, e.g., Apple. Retrieval-induced forgetting predicts that the partial activation and subsequent suppression of Fruit–Apple in favor of Fruit–Pear will lead to worse memory for Fruit–Apple. In fact, Anderson and Spellman found that subjects' memory for Fruit–Apple was weakened when compared with witnessed pairs that had not been partially activated, such as Sport–Tennis. The suppression only held for pairs such as Fruit–Apple that involved prototypical exemplars of the superordinate category (here, Fruit), because non-prototypical exemplars are less strongly associated with the category. As expected, then, memory for Fruit–Kiwi was not weakened.

https://doi.org/10.1017/langcog.2016.17 Published online by Cambridge University Press

Retrieval-induced forgetting predicts that a construction that is in competition will be weakened whenever another form 'wins' (is used). For example, if, whenever a double-object pattern with *explain*, as in (35), is expected, (36) is repeatedly and consistently witnessed instead, (35) will become harder to retrieve. In this way, (36) will come to preempt (35).

(35) ??She explained him something.
(36) She explained something to him.

3.2. PREDICTIONS AS CONDITIONAL PROBABILITIES

As explained in Goldberg (2011a), the probability of a construction CxB statistically preempting CxA for a particular verb, $verb_i$, is:

(37) P(CxB | context suitable for CxA, and $verb_{i.}$)

For example, if we assume that *explain* does not readily occur in the double-object construction because it is statistically preempted by the *to*-dative construction, we predict the probability in (38) to be high:

(38) P(dative | context suitable for the double-object construction and *explain*)

In order to operationalize how to count 'contexts that are at least as suitable for the double-object construction', we can use the total number of double-object and *to*-dative uses in a given corpus, when the semantics and information structure of the double-object construction are satisfied. That is,

(39) P(dative | context suitable for double-object construction and $verb_{i.}$) ≈ P(dative | $verb_{i.}$ and (dative$_{\text{with relevant restrictions}}$ *or* double-object construction))

In fact, this probability has been estimated to be quite high (.99) on the basis of a corpus analysis (Goldberg, 2011a).

Also relevant is the frequency with which the preempting situation is witnessed. That is, suppose that the first time a learner hears *explain*, she expects to hear it used in the double-object construction, but instead hears it used in the *to*-dative. At that moment, the probability of witnessing *explain* in a preemptive context is 1, but only a single case has been witnessed. Clearly, the learner should not infer from a single exposure that the double-object construction is preempted for *explain*. On the other hand, if a learner hears *explain* used datively 100 times, without ever hearing it used in the double-object construction, the probability hasn't changed – it is still 1 – but the CONFIDENCE of preemption should be increased. In fact, it has been demonstrated experimentally that essentially a gap is more likely to be considered to be non-accidental when the overall token frequency is increased (Reeder, Newport, & Aslin, 2013; Xu & Tenenbaum, 2007). We can observe further that it is not

https://doi.org/10.1017/langcog.2016.17 Published online by Cambridge University Press

likely that confidence increases linearly with frequency, so we appeal to the logarithmic function. Thus we can separate the two factors that determine the strength of preemption as follows: Probability (40), and Confidence (41):

(40) *Probability* of CxB statistically preempting CxA for $verb_i$:
P(CxB| contexts in which CxA would be suitable)

(41) *Confidence* of statistical preemption for $verb_i$, where F=frequency:
ln F(CxB when CxA would be suitable)

4. Conclusion

Constructions are typically partially but not fully productive. The present paper sketches the two complementary factors: DYNAMIC CATEGORIZATION and STATISTICAL PREEMPTION. Much more work is needed to provide a fully comprehensive and explicit account (see Goldberg & Ambridge, forthcoming), but it is clear that, as learners record statistics of their language, they dynamically categorize their input on the basis of form and function. Productivity is to a large extent determined by coverage, which is a general principle of induction: essentially, a potential new coinage is judged acceptable to the extent that the formal linguistic category it would join is well attested by similar exemplars. This idea captures the fact that each construction has a restricted range of distribution, typically dependent on various semantic, pragmatic, and phonological properties of the exemplars that are witnessed.

Recognizing that categories do not exist in isolation from one another, it is also important to recognize a process of statistical preemption whereby learners learn to avoid using one construction, even when the construction's constraints would seem to be satisfied, if an alternative formulation has been systematically witnessed instead. The mechanism required for statistical preemption is competition-driven learning, which is a domain-general process. When two competitors are activated, but one systematically wins, the loser becomes less accessible over time. In this way, with a recognition of both general properties of categorization and the role of competition among categories, we can begin to EXPLAIN OURSELVES THE PARADOX OF PARTIAL PRODUCTIVITY.

REFERENCES

Ambridge, B., Pine, J. M., & Lieven, E. V. (2015). Explanatory adequacy is not enough: response to commentators on 'Child language acquisition: why universal grammar doesn't help'. *Language*, **91**(3), e116–e126.

Ambridge, B., Pine, J. M., & Rowland, C. F. (2012). Semantics versus statistics in the retreat from locative overgeneralization errors. *Cognition*, **123**(2), 260–279.

Ambridge, B., Pine, J. M., Rowland, C. F., & Chang, F. (2012). The roles of verb semantics, entrenchment and morphophonology in the retreat from dative argument structure overgeneralization errors. *Language*, **88**(1), 45–81.

https://doi.org/10.1017/langcog.2016.17 Published online by Cambridge University Press

Ambridge, B., Pine, J. M., Rowland, C. F., Jones, R. L., & Clark, V. (2009). A semantics-based approach to the 'no negative evidence' problem. *Cognitive Science*, **33**(7), 1301–1316.

Ambridge, B., Pine, J. M., Rowland, C. F., & Young, C. R. (2008). The effect of verb semantic class and verb frequency (entrenchment) on children's and adults' graded judgements of argument-structure overgeneralization errors. *Cognition*, **106**(1), 87–129.

Anderson, M. C., Green, C., & McCulloch, K. C. (2000). Similarity and inhibition in long-term memory: evidence for a two-factor theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **26**(5), 1141–1159.

Anderson, M. C., & Spellman, B. A. (1995). On the status of inhibitory mechanisms in cognition: memory retrieval as a model case. *Psychological Review*, **102**(1), 68–100.

Arnold, J. E., Eisenband, J. G., Brown-Schmidt, S., & Trueswell, J. C. (2000). The rapid use of gender information: evidence of the time course of pronoun resolution from eyetracking. *Cognition*, **76**(1), 13–26.

Aronoff, M. (1976). *Word formation in generative grammar* (Linguistic Inquiry Monograph 1). Cambridge, MA: MIT Press.

Artola, A., Brocher, S., & Singer, W. (1990). Different voltage-dependent thresholds for inducing long-term depression and long-term potentiation in slices of rat visual cortex. *Nature*, **347**, 69–72.

Baker, C. L. (1979). Syntactic theory and the projection problem. *Linguistic Inquiry*, **10**(4), 533–581.

Barðdal, J. (2008). *Productivity: evidence from case and argument structure in Icelandic*. Amsterdam: John Benjamins.

Bolinger, D. (1977). *Meaning and form*. London: Longman.

Bowerman, M. (1988). The 'no negative evidence' problem: How do children avoid constructing an overly general grammar? In J. Hawkins (Ed.), *Explaining language universals* (pp. 73–101). Oxford: Basil Blackwell.

Bowerman, M. (1996). Argument structure and learnability: Is a solution in sight? In J. Johnson, M. L. Juge, & J. L. Moxley (Eds.), *Proceedings of the twenty-second annual meeting of the Berkeley Linguistics Society: general session and parasession on the role of learnability in grammatical theory* (pp. 454–468). Berkeley, CA: Berkeley Linguistics Society.

Bowerman, M., & Choi, S. (2001). Shaping meaning for language: universal and language-specific in the acquisition of spatial semantic categories. In M. Bowerman and S. C. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 475–511). Cambridge: Cambridge University Press.

Bowerman, M., & Choi, S. (2003). Space under construction: language-specific spatial categorization in first language acquisition. In *Language in mind: advances in the study of language and cognition* (pp. 387–428). Cambridge, MA: MIT Press.

Boyd, J. K., & Goldberg, A. E. (2009). Input effects within a constructionist framework. *Modern Language Journal*, **93**(iii), 418–429.

Boyd, J. K., & Goldberg, A. E. (2011). Learning what not to say: the role of statistical preemption and categorization in a-adjective production. *Language*, **87**(1), 55–83.

Braine, M. D. S. (1971). The acquisition of language in infant and child. In C. Reed (Ed.), *The learning of language* (pp. 7–95). New York: Appleton-Century-Crofts.

Braine, M. D. S., & Brooks, P. J. (1995). Verb argument structure and the problem of avoiding an overgeneral grammar. In M. Tomasello & W. E. Merriman (Eds.), *Beyond names for things: young children's acquisition of verbs* (pp. 353–376). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Bresnan, J., Cueni, A., Nikitina, T., & Baayen, R. H. (2007). Predicting the dative alternation. In G. Bouma, I. Kraemer, & J. Zwarts (Eds.), *Cognitive foundations for interpretation* (pp. 69–94). Amsterdam: Royal Netherlands Academy of Science.

Bresnan, J., & Ford, M. (2010). Predicting syntax: processing dative constructions in American and Australian varieties of English. *Language*, **86**(1), 186–213.

Bresnan, J., & Hay, J. (2008). Gradient grammar: an effect of animacy on the syntax of *give* in New Zealand and American English. *Lingua*, **118**(2), 245–259.

Brooks, P. J., & Tomasello, M. (1999). How children constrain their argument structure constructions. *Language*, **75**(4), 720–738.

https://doi.org/10.1017/langcog.2016.17 Published online by Cambridge University Press

Brooks, P. J., and Zizak, O. (2002) Does preemption help children learn verb transitivity? *Journal of Child Language*, **29**(4), 759–781.

Brown, R., & Hanlon, C. (1970). Derivational complexity and order of acquisition in child speech. In J. R. Hayes (Ed.), *Cognition and the development of language*. New York: Wiley.

Bybee, J. (1985). *Morphology: a study of the relation between meaning and form*. Amsterdam/Philadelphia, PA: John Benjamins Publishing.

Bybee, J. (1995). Regular morphology and the lexicon. *Language and Cognitive Processes*, **10**(5), 425–455.

Bybee, J. (2010) *Language, usage and cognition*. Cambridge: Cambridge University Press.

Bybee, J., & Eddington, D. (2006). A usage-based approach to Spanish verbs of 'becoming'. *Language*, **82**(2), 323–355.

Casenhiser, D., & Goldberg, A. E. (2005). Fast mapping between a phrasal form and meaning. *Developmental Science*, **8**(6), 500–508.

Clark, E. V. (1987). The principle of contrast: a constraint on language acquisition. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 1–33). Hillsdale, NJ: Lawrence Erlbaum Associates.

Clausner, T. C., & Croft, W. (1997). Productivity and schematicity in metaphors. *Cognitive Science*, **21**(3), 247–282.

Collins, P. (1995). The indirect object construction in English: an informational approach. *Linguistics*, **33**(1), 35–49.

Coppock, E. (2008). *The logical and empirical foundations of Baker's paradox*. Stanford, CA: Stanford University Press.

Croft, W., & Cruse, D. A. (2004). *Cognitive linguistics*. Cambridge: Cambridge University Press.

Davies, M. (2008a). The Corpus of Contemporary American English: 450 million words, 1990–present. Online: <http://corpus.byu.edu/coca/>.

Davies, M. (2008b). The Corpus of Historical American English: 400 million words, 1810–2009. Online: <http://corpus.byu.edu/coha/>.

Desagulier, G. (2015). A lesson from associative learning: asymmetry and productivity in multiple-slot constructions. *Corpus Linguistics and Linguistic Theory*, online: <doi:10.1515/cllt-2015-0012>.

Dryer, M. S. (1986). Primary objects, secondary objects, and antidative. *Language*, **62**, 808–845.

Ellis, N. C., & Ferreira-Junior, F. (2009). Constructions and their acquisition: islands and the distinctiveness of their occupancy. *Annual Review of Cognitive Linguistics*, **7**(1), 187–220.

Erteschik-Shir, N. (1979). Discourse constraints on dative movement. In S. Laberge & G. Sankoff (Eds.), *Syntax and semantics* (pp. 441–467). New York: Academic Press.

Foraker, Stephani, Regier, Terry, Khetarpal, Naveen, Perfors, Amy, & Tenenbaum, Joshua B. (2007). Indirect evidence and the poverty of the stimulus: the case of anaphoric *one*. In D. S. McNamara & J. G. Trafton (Eds.), *Proceedings of the twenty-ninth annual conference of the Cognitive Science Society* (pp. 275–280). New York: Lawrence Erlbaum.

Givón, T. (1979). *On understanding grammar*. New York: Academic Press.

Givón, T. (1984). *Syntax: A functional-typological introduction*. Amsterdam: Benjamins.

Goldberg, A. E. (1993). Another look at some learnability paradoxes. Paper presented to the Proceedings of the 25th Annual Stanford Child Language Research Forum, Stanford, 1993.

Goldberg, A. E. (1995). *Constructions: a construction grammar approach to argument structure*. Chicago, IL: Chicago University Press.

Goldberg, A. E. (2006). *Constructions at work: the nature of generalization in language*. Oxford: Oxford University Press.

Goldberg, A. E. (2011a). Corpus evidence of the viability of statistical preemption. *Cognitive Linguistics*, **22**(1), 131–153.

Goldberg, A. E. (2011b). Are a-adjectives like afraid prepositional phrases underlying and does it matter from a learnability perspective? Unpublished ms, Princeton University.

Goldberg, A. E., and Ambridge, Ben (forthcoming). *Explain me this*. Princeton, NJ: Princeton University Press.

Goldberg, A. E., & Boyd, J. K. (2015). *A*-adjectives, statistical preemption, and the evidence: reply to Yang (2015). *Language*, e184–e197.

https://doi.org/10.1017/langcog.2016.17 Published online by Cambridge University Press

Goldberg, A. E., Casenhiser, D., & Sethuraman, N. (2004). Learning argument structure generalizations. *Cognitive Linguistics*, **15**, 289–316.

Green, G. M. (1974). *Semantics and syntactic regularity*. Bloomington, IN: Indiana University Press.

Gropen, J., Pinker, S., Hollander, M., & Goldberg, R. (1991). Syntax and semantics in the acquisition of locative verbs. *Journal of Child Language*, **18**, 115–151.

Gropen, J., Pinker, S., Hollander, M., Goldberg, R., & Wilson, R. (1989). The learnability and acquisition of the dative alternation in English. *Language*, **65**(2), 203–257.

Hilpert, M. (2013). *Constructional change in English: developments in allomorphy, word formation, and syntax*. Cambridge: Cambridge University Press.

Janda, R. D. (1990). Frequency, markedness and morphological change: on predicting the spread of noun-plural -s in Modern High Germanic and West Germanic. *Proceedings of the 7th Eastern States Conference on Linguistics*, 136–153. Online: <http://files.eric.ed.gov/fulltext/ED333749.pdf#page=145>.

Johnson, M. A., Turk-Browne, N., & Goldberg, A. E. (2013). Prediction plays a key role in language development as well as processing. *Behavioral and Brain Sciences*, **36**(4), 360–361.

Kalyan, S. (2012). Similarity in linguistic categorization: the importance of necessary properties. *Cognitive Linguistics*, **23**(3), 539–554.

Kim, G., Lewis-Peacock, J. A., Norman, K. A., & Turk-Browne, N. B. (2014). Pruning of memories by context-based prediction error. *Proceedings of the National Academy of Sciences*, **111**(24), 8997–9002.

Kiparsky, P. (1993). Blocking in non-derived environments. In S. Hargus & Ellen Kaisse (Eds.), *Phonetics and Phonology 4: Studies in Lexical Phonology* (pp. 277–313). San Diego, CA: Academic Press.

Kroll, J. F., & Stewart, E. (1994). Category inference in translation and picture naming: evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, **33**, 149–174.

Kuhl, P. K. (2000). A new view of language acquisition. *Proceedings of the National Academy of Sciences*, **97**(22), 11850–1857.

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, **307**, 161–163.

Lakoff, G. (1987). *Women, fire and dangerous things*. Chicago, IL: University of Chicago Press.

Landauer, T. K. (2006). Latent semantic analysis. *Encyclopedia of Cognitive Science*. Online: <http://onlinelibrary.wiley.com/doi/10.1002/0470018860.s00561/abstract>.

Langacker, R. W. (1987). *Foundations of cognitive grammar: theoretical prerequisites* (Vol. 1). Stanford, CA: Stanford University Press.

Madlener, Karin. (2015). *Frequency effects in instructed second language acquisition*. Berlin: Mouton de Gruyter.

Marcotte, J. (2005). *Causative alternation errors in child language acquisition*. Unpublished PhD thesis, Stanford.

Marcus, G. F. (1993). Negative evidence in language acquisition. *Cognition*, **46**, 53–85.

McDonough, K., & Nekrasova-Becker, T. (2014). Comparing the effect of skewed and balanced input on English as a foreign language learners' comprehension of the double-object dative construction. *Applied Psycholinguistics*, **35**(2), 419–442.

McRae, K., Spivey-Knowlton, M. J., & Tanenhaus, M. K. (1998). Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language*, **38**, 283–312.

Newman, E. L., & Norman, K. A. (2010). Moderate excitation leads to weakening of perceptual representations. *Cerebral Cortex*, **20**(11), 2760–2770.

Norman, K., Newman, E. L., & Detre, G. (2007). A neural network model of retrieval-induced forgetting. *Psychological Review*, **114**(4), 887–953.

Oehrle, R. (1975). *The grammatical status of the English dative alternation*. Unpublished PhD dissertation, MIT.

Osherson, D. N., Smith, E. E., Wilkie, O., Lopez, A., & Shafir, E. (1990). Category-based induction. *Psychological Review*, **97**(2), 185–200.

https://doi.org/10.1017/langcog.2016.17 Published online by Cambridge University Press

Perek, F. (2016). Using distributional semantics to study syntactic productivity in diachrony: a case study. *Linguistics*, **54**(1), 149–188.
Perek, F., & Goldberg, A. E. (2015). Generalizing beyond the input: the functions of the constructions matter. *Journal of Memory and Language*, **84**, 108–127.
Pickering, M. J., & Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences*, **11**(3), 105–110.
Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, **36**(4), 329–347.
Pinker, S. (1989). *Learnability and cognition: the acquisition of argument structure*. Cambridge, MA: MIT Press.
Rainer, S. (1988). A short story of down. In W. Hüllen, & R. Schulze (Eds.), *Understanding the lexicon: meaning sense and world knowledge in lexical semantics* (pp. 394–410). Tübingen: Max Niemeyer Verlag.
Reeder, P, Newport, Elissa, & Aslin, Richard N. (2013). From shared contexts to syntactic categories: the role of distributional information in learning linguistic form-classes. *Cognitive Psychology*, **66**(1), 30–54.
Robenalt, C., & Goldberg, A. E. (2015). Judgment and frequency evidence for statistical preemption: it is relatively better to *vanish* than to *disappear* a rabbit, but a lifeguard can equally well *backstroke* or *swim* children to shore. *Cognitive Linguistics*, **26**(3), 467–503.
Robenalt, C., & Goldberg, A. E. (2016). Nonnative speakers do not take competing alternative expressions into account the way native speakers do. *Language Learning*, **66**(1), 60–93.
Stefanowitsch, A. (2008). Negative entrenchment: a usage-based approach to negative evidence. *Cognitive Linguistics*, **19**(3), 513–531.
Stephens, G. J., Silbert, L. J., & Hasson, U. (2010). Speaker–listener neural coupling underlies successful communication. *Proceedings of the National Academy of Sciences*, **107**(32), 14425–14430.
Storm, B. C., & Levy, B. J. (2012). A progress report on the inhibitory account of retrieval-induced forgetting. *Memory & Cognition*, **40**(6), 827–843.
Suttle, L., & Goldberg, A. E. (2011). The partial productivity of constructions as induction. *Linguistics*, **49**(6), 1237–1269.
Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, **268**(5217), 1632–1634.
Taylor, J. R. (2003). *Linguistic categorization*. Oxford: Oxford University Press.
Theakston, A. L. (2004). The role of entrenchment in constraining children's verb argument structure overgeneralisations: a grammatical judgment study. *Cognitive Development*, **19**, 15–34.
Thompson, S. A. (1990). Information flow and 'dative shift' in English. In J. Edmondson, K. Feagin, & P. Mühlhäusler (Eds.), *Development and diversity: linguistic variation across time and space* (pp. 239–253). Dallas, TX: Summer Institute of Linguistics.
Thompson, S. A. (1995). The iconicity of 'dative shift' in English: considerations from information flow in discourse. In M. E. Landsberg (Ed.), *Syntactic iconicity and linguistic freezes: the human dimension* (Studies in Anthropological Linguistics 9) (pp. 155–175). Berlin / New York: Mouton de Gruyter.
Tomasello, M. (2003). *Constructing a language: a usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
Wasow, T. (2002). *Postverbal behavior*. Stanford, CA: CSLI.
Wonnacott, E., Boyd, J. K., Thompson, J., & Goldberg, A. E. (2012). Novel construction learning in five year olds. *Journal of Memory and Language*, **66**, 458–478.
Xu, F., & Tenenbaum, J. B. (2007). Word learning as Bayesian inference. *Psychological Review*, **114**(2), 245–272.
Yang, Charles (2015). Negative knowledge from positive evidence. *Language*, **91**(4), 938–953.
Zeschel, A., & Bildhauer, F. (2009). Islands of acceptability. Paper presented at the AfLiCO, Paris.

https://doi.org/10.1017/langcog.2016.17 Published online by Cambridge University Press