Input Effects Within a Constructionist Framework

JEREMY K. BOYD
Princeton University
Department of Psychology
Green Hall
Princeton, NJ 08544
Email: jkboyd@princeton.edu

ADELE E. GOLDBERG
Princeton University
Department of Psychology
Green Hall
Princeton, NJ 08544
Email: adele@princeton.edu

Constructionist approaches to language hypothesize that grammar can be learned from the input using domain-general mechanisms. This emphasis has engendered a great deal of research—exemplified in the present issue—that seeks to illuminate the ways in which input-related factors can both drive and constrain constructional acquisition. In this commentary piece, we situate results reported by contributors to the present issue within the larger body of acquisition work in the constructionist framework. We address the importance of both type frequency and skewed input samples in the development of constructional categories and we compare different ways that the association between verbs and constructions can be measured, including through the use of conditional probabilities, lexical biases, and introspective judgments.

THE CENTRAL TENET OF CONSTRUCTIONIST approaches to language is that grammatical knowledge is instantiated in learned form–function pairings known as *constructions* (Culicover & Jackendoff, 2005; Fillmore, Kay, & O'Connor, 1988; Goldberg, 1995, 2006; Kay & Fillmore, 1999; Tomasello, 2003). Constructions vary independently along two dimensions: size and abstractness. This entails that constructions can be both small and large (e.g., both words and multiword phrasal patterns are constructions) and that they can be both concrete and abstract. For example, the concrete *dog* and the abstract N are both smaller constructions; likewise, the concrete imperative *Ask not what your country can do for you*... and the abstract passive Subj Aux VP (PP_{by}) are both larger constructions. In short, although they clearly vary in terms of their size and degree of internal complexity, *all* linguistic symbols are constructions. This feature of the theory stands in opposition to Chomskyan approaches to language (e.g., Chomsky, 1995), which view constructions as the epiphenomenal product of syntactic operations (grammar) working over a concrete lexicon that includes only morphemes, words, and idiomatic expressions.

The Chomskyan paradigm has come under increasing criticism in recent years for its claim that grammar cannot be learned from the input and therefore only emerges during the course of acquisition thanks to the contribution of innate, language-specific constraints (Christiansen & Chater, 2008; Deacon, 1997; Elman et al., 1996; Langacker, 1987; Lewis & Elman, 2001; Perfors, Tenenbaum, & Regier, 2006; Ramscar & Yarlett, 2007; Tomasello, 2003). This view is in part motivated by the overly complex character of Chomskyan grammatical models. Because they appeal to underlying levels of representation and to invisible elements, it is difficult to understand how acquisition could proceed solely from the input. As a means of overcoming this hurdle, it is assumed that domain-specific biases are required during development. In contrast, constructionist models sidestep nativist logic by embracing a simpler notion of syntax (e.g., Culicover & Jackendoff, 2005). According to the constructionist approach, meanings are directly encoded in surface-level forms, and no movement, empty, null, or silent syntactic elements of any kind are posited. This reduction in underlying complexity

The Modern Language Journal, 93, iii, (2009)
0026-7902/09/418–429 $1.50/0

leads to a concomitant reduction in the number of innate, domain-specific stipulations that are required to make the model learnable (Goldberg, 2006; Tomasello, 2000, 2003). In fact, the constructionist approach takes it as a null hypothesis that natural languages can be learned without reference to *any* innate, specifically linguistic principles.

The assumption that linguistic knowledge is learned rather than innately given drives a research program that emphasizes the use of domain-general mechanisms to learn both first language (L1) and second language (L2) grammars; it also encourages detailed investigations into the nature of the input. The articles in the present issue fall squarely within this tradition. In particular, they focus on L2 acquisition and offer a wide array of new data indicating that the input often has surprising characteristics that can both constrain and drive learning.

The following commentary is organized into five sections. In the first section we review new evidence from Collins, Trofimovich, White, Cardoso, and Horst (this issue) and McDonough and Kim (this issue) indicating that perceptual salience and type frequency are among the factors that have a significant effect on constructional acquisition. In the second section we turn to a less widely discussed phenomenon—the facilitatory effect of skewed input on construction learning—and discuss its theoretical underpinnings, as well as the experimental studies that have documented it. In the third section we review results from Ellis and Ferreira-Junior (this issue) and Wulff, Ellis, Römer, Bardovi-Harlig, and LeBlanc (this issue) indicating that skewed constructional input is widespread across many types of constructions. We also compare different ways in which the association between verbs and constructions can be measured, including the use of conditional probabilities, lexical biases (as determined by multiple distinctive collexeme analysis [MDCA]; Wulff et al.), and introspective judgments (Ellis & Ferreira-Junior). The fourth section then addresses an experimental study by Year and Gordon (this issue) that finds a consistent null effect of skewed input on Korean speakers' acquisition of the English ditransitive construction. We highlight the fact, however, that when the experiment's full data set is considered (see Year, 2009), there is evidence that suggests that skewed input gives participants an increased awareness of semantic constraints on the use of the ditransitive early on. In the fifth section we conclude by foregrounding the connection between the empirical findings disseminated in the present issue and the constructionist approach's commitment to developing a model of language acquisition that emphasizes the use of domain-general mechanisms to learn grammar from richly structured input.

PERCEPTUAL SALIENCE AND TYPE FREQUENCY

It is a truism that some things are more difficult to learn than others. In this respect, language acquisition is no different from the development of other complex behaviors. The language learner is confronted by an array of form–function pairings that must be acquired, not all of which are equally amenable to analysis. One question that this state of affairs poses for language researchers is *why* it is that some things are more difficult to learn than others.

Collins et al. (this issue) provide a number of answers to this question. First, they note that language development in hearing populations begins with an auditory signal that must be decoded. Not every signal, however, is equally clear. Easy-to-learn and difficult-to-learn constructions can, for example, sometimes be distinguished from one another by appealing to relative differences in perceptual salience, among other features (Slobin, 1982); that is, one of the reasons that some constructions are more difficult for learners to acquire is that they are simply *less prominent* in the speech stream. Collins et al. show, for instance, that the past tense suffix is more difficult to hear because it is rarely stressed or followed by a pause, only occurs as a separate syllable in its [–əd] allomorph, and is often minimally articulated. In contrast, the progressive marker is significantly easier to hear: It always receives stress, is sometimes followed by a pause, and always occurs as an intact syllable. The perceptual salience that these two markers have in the speech stream thus plays a potentially significant role in how easily they are acquired.

Second, type frequency undoubtedly affects what learners find more or less difficult to learn (Bybee, 1985; Plunkett & Marchman, 1991, 1993). All else being equal, constructions that occur earlier in the developmental sequence may do so because learners have more opportunities to see them instantiated with many different lexical items. Returning to a comparison of the simple past and progressive constructions, Collins et al. (this issue) argue that only four verb types from the thousand most frequent English word families commonly occur in the input with the regular (*–ed*) past tense marker: *ask*, *happen*, *change*, and *decide*. In stark contrast, the progressive marker (–*ing*) commonly occurs with at least 26 different verb types. The implication here is that early acquisition of the progressive construction is facilitated by relatively high type frequency in the input.

Further evidence of the association between type frequency and ease of acquisition is provided in an experimental study by McDonough and Kim (this issue). The authors asked whether native speakers of Thai who were learning English were better at forming English questions when they had participated in a structural priming task that featured exposure to many lexical verbs or few lexical verbs. Their results show a reliable facilitatory effect of high type frequency on question production. Specifically, when participant utterances in the priming task were prompted with 36 different verbs rather than just 6, their production performance on a posttest was significantly better. This outcome has important ramifications for the significance of high-variance input at later stages of L2 acquisition, a topic that will be addressed in more detail in the fourth section.

The results from McDonough and Kim (this issue) and Collins et al. (this issue) are mutually supportive and are in accordance with numerous findings from other researchers who have emphasized the relationship between type frequency and the development of the abstract representations needed to support productivity (Boyd, 2007; Bybee, 1985, 1995; Ninio, 2005; Onnis, Waterfall, & Edelman, 2008; Plunkett & Marchman, 1991, 1993; Tomasello, 2003). As type frequency goes up, it becomes increasingly clear to learners that existing item-based constructional schemas can be generalized. This leads to the development of progressively more abstract representations, which can be deployed to produce and understand utterances that were not present in the input.

High type frequency often goes hand in hand with high variability among the exemplars that make up a constructional category. The more verbs that occur in an argument structure construction, for example, the more perceptually variable the construction will be to learners. There is thus a clear relationship between type frequency and variability, one that seems to favor increased variability during acquisition as a means of elaborating on item-specific constructional schemas.

In the remainder of this commentary, however, we focus on a seemingly contradictory hypothesis: that a particular type of low-variance input, *skewed input,* also has a facilitatory effect on constructional acquisition.

LOW–VARIANCE INPUT AND CONSTRUCTION LEARNING

Constructional acquisition can be viewed essentially as a process of categorization (Goldberg, 2006; Goldberg, Casenhiser, & White, 2007; Wulff et al., this issue). Children's linguistic competence can be said to include knowledge of abstract constructions when they are able to implicitly recognize that structurally identical sentences populated by different sets of lexical items nevertheless have the same basic meaning (Bencini & Valian, 2008; Huttenlocher, Vasilyeva, & Shimpi, 2004; Savage, Lieven, Theakston, & Tomasello, 2003, 2006). This entails, for example, that when individuals hear the sentences *John gave Mary the keys* and *The accountant sent his client an invoice*, they know that these items are similar in both form (they share the same Subj V Obj_1 Obj_2 structure) and meaning (they refer to events in which an agent transfers an item to a recipient). Once abstract constructional categories are in place, they can be used to produce novel utterances and to interpret sentences that have never been heard previously.

A number of studies on nonlinguistic categorization have shown that the level of variability present in learners' input affects their ability to form new categories (Casasola, 2005; Elio & Anderson, 1984; Gentner, Loewenstein, & Hung, 2007; Gentner & Medina, 1998; Mervis & Rosch, 1981; Posner & Keele, 1968). Specifically, participants are better able to induce new categories if they are initially exposed to a low-variance sample rather than a high-variance sample. As an example of this principle, consider a child whose task it is to learn the general concept "bird" based on either the initial input stimuli represented by the top row of Figure 1—a relatively low-variance sample—or the input stimuli represented by the bottom row of Figure 1—a relatively high-variance sample.

Exposure to the birds in the top row will engender faster category development because they can be classified together based on readily apparent similarities in size, behavior, and habitat. Subsequent generalization of the category to new members is thus facilitated by having an initial input sample that makes grouping more obvious. However, early exposure to the full range of instances (cf. the birds in the bottom panel) obscures commonalities and retards initial category formation. It is simply less obvious that penguins, robins, and ostriches are the same sorts of things, because they differ significantly on a number of dimensions, including size, habitat, and means of locomotion.

FIGURE 1
Examples of Input Stimuli for the Concept "Bird."

Note. Categories (e.g., BIRD) are easier to acquire when the items to be categorized constitute a low-variance sample (top row) than a high-variance sample (bottom row).

Initial exposure to a low-variance sample has been found to facilitate the categorization of shapes (Posner & Keele, 1968), spatial relations (Casasola, 2005), and people into social groups (Elio & Anderson, 1984). More importantly for the present purposes, low-variance input has been shown to play a facilitatory role in the initial induction of constructional categories in language (Casenhiser & Goldberg, 2005; Maguire, Hirsh-Pasek, Golinkoff, & Brandone, 2008). Because constructional slots like V and NP can be realized in many different ways, abstract semantic and structural similarities across two exemplars of the same construction are not as readily apparent to learners as they could be (Gentner & Medina, 1998). Fortunately, however, many aspects of the linguistic input appear to be naturally structured so as to reduce this variability.

Zipf's Law (1935) states that, given a corpus of natural language utterances, the frequency of any word in the corpus will be inversely proportional to its rank in a frequency table. The fact that words are distributed in Zipfian fashion has tremendous repercussions on input variability. It means that although individuals may have upward of 60,000 words in their vocabulary, the number of words that they use with any regularity will be at least an order of magnitude smaller. This constitutes a significant decrease in input variability: Learners tend to hear the same set of high-frequency, high-utility words over and over again (Mintz, Newport, & Bever, 2002). Consequently, these words will be learned more quickly than if each word in the language had the same probability of occurrence.

Goldberg, Casenhiser, and Sethuraman (2004) noted that just as words have a Zipfian distribution in a language, so too can verbs in particular constructions (see also Ninio, 1999). In a corpus study, Goldberg et al. demonstrated that for three construction types—the intransitive motion construction (e.g., *The fly buzzed into the room*), the caused-motion construction (e.g., *Pat sneezed the* *foam off of the cappuccino*), and the ditransitive construction (e.g., *She faxed him a letter*)—the input that children received from their primary caregivers was centered around a handful of general-purpose verbs (Clark, 1996). In particular, *go*, *put*, and *give* were by far the most commonly used verbs in the three constructions, with *go* occurring in 39% of all instances of the intransitive motion construction, *put* occurring in 38% of all instances of the caused-motion construction, and *give* occurring in 20% of all instances of the ditransitive construction. Further, an analysis of child productions in the same corpus showed that children's patterns of usage strongly mirrored those of their primary caregivers.

This type of distribution has a potentially significant impact on constructional development because it means that these constructional categories have obvious, verb-centered prototypes and that the overall variability among all of the exemplars of a particular construction type is not nearly as high as it would be if a large number of verbs occurred equally often in the construction. The hypothesis that follows is that the presence of low-variance, Zipfian input facilitates construction learning by making both the meaning and form of a particular construction simpler to identify.

Evidence supporting this hypothesis comes from a series of experiments with children (Casenhiser & Goldberg, 2005) and adults (Goldberg et al., 2004) in which participants received brief exposure to a novel construction and were then tested on their ability to comprehend never-before-seen exemplars of the construction. Overall type and token frequency were controlled for, so all subjects saw the same number of exemplars involving the same number of (nonsense) verbs. In the *skewed* (low-variance) input condition, however, one verb accounted for half of the input, whereas in the *balanced* frequency condition, the verbs were more evenly distributed. In this way, learners with low-variance input did not have lower type frequency (or token frequency); rather, the skewed input was low-variance in the sense that there was more repetition of a particular verb. The results demonstrate that participants whose exposure had consisted of a skewed input sample performed significantly better at generalizing to new items at test than those whose exposure had been balanced. A similar outcome is reported in Maguire et al. (2008): Skewed input also facilitates the acquisition of novel verbs.

To understand how skewed input might have this sort of effect on construction learning, consider the case of the caused-motion construction, which is disproportionally realized with *put* (Goldberg et al., 2004). The meaning of *put* is highly correlated with that of the construction, which denotes events in which an agent moves an item to (or from) a location. This allows for *put* to anchor children's initial conception of the semantics of the caused-motion construction (Goldberg, 2006). Moreover, the fact that the input is centered around *put* reduces formal variability among the different caused-motion exemplars that children are exposed to, which makes the abstract structural properties of the construction more easily resolvable (Gentner & Medina, 1998; Gentner et al., 2007). The V slot in the construction, for instance, is simpler for learners to recognize, to the extent that it is consistently realized with the same verb.

NEW RESULTS AND INTERPRETATIONS RELATING TO SKEWED INPUT

A number of the articles in the present issue expand on the pervasiveness and significance of skewed input for construction learning. In particular, Wulff et al. (this issue) and Ellis and Ferreira-Junior (this issue) provide detailed corpus analyses indicating that verbs have Zipfian distributions in a wide variety of constructions—both as measured in terms of the exemplars that L2 learners hear and in the utterances that they produce. This sort of work demonstrates that skewed frequencies are quite general: Specific verbs seem to be prominently featured in many constructions (Goldberg, 1996; Goldberg et al., 2004; Kidd, Lieven, & Tomasello, 2006; Sethuraman, 2002; Thompson, 2002). Ellis and Ferreira-Junior make the further point that the highest frequency verb tends to account for an even greater share of the overall tokens when there are fewer types per construction, as is the case for early learners.

However, what does it mean to say that the distribution of verbs is skewed with respect to a particular construction? How do we measure verb–construction associations? There are, it turns out, a number of different approaches. Wulff et al. (this issue) argue that the proper measure is to consider a verb's bias toward one construction versus others. In contrast, Ellis and Ferreira-Junior (this issue) attempt to measure verb–construction associations by asking participants to explicitly rate the strength of the semantic overlap between verb meaning and construction meaning. In the text that follows, we juxtapose these two alternatives with the use of conditional probabilities as a measure of skewing.

Goldberg and colleagues determined whether a construction has skewed token frequency by considering the conditional probability that a verb occurs in a given construction: $P(\text{verb}_i \mid \text{construction}_k)$ (Goldberg, 2006; Goldberg et al., 2004). For example, *make* accounts for 20% of the tokens of the *way* construction (e.g., *She made her way into the room*). This is far more than any other single verb (Goldberg, 1996), so tokens of the construction are skewed toward *make*: $P(\textit{make} \mid \textit{way}$ construction$) = .20$; for all other verbs, $P(\text{verb}_{i \neq \text{make}} \mid \textit{way}$ construction$) < .05$).

The fact that *make* is frequent overall in the language with a range of different senses is not relevant to this measure; neither is the fact that the construction is relatively rare in the language overall. Whether *make* happens to be the very first verb uttered in the construction is also not relevant: More than one factor may influence which words happen to be uttered first (Goldberg, 2006; Tomasello, 2003; but see Ellis & Ferreira-Junior, this issue, for an argument that the most frequent verbs are the earliest learned; also Ninio, 1999). What *is* relevant to determining whether the frequency of an argument structure construction (k) is skewed is the set of conditional probabilities for all n verbs that appear in k: For all verbs, $i = 1$ to n, $P(\text{verb}_i \mid \text{construction}_k)$.

Many constructions (but not all—see below) display a skewed distribution in this way for the following reason. Certain verbs occur much more often in their constructions because they have a wider range of functions than their competitors (Ellis & Ferreira-Junior, this issue; Goldberg, 2006; Goldberg et al., 2004). *Go*, for instance, has a much more general meaning than either *walk* or *run*. This means that it can be used in the same situations that *walk* and *run* are, and countless others. Small numbers of verbs thus account for large numbers of tokens of a construction because they are pragmatically useful and can describe many different types of events.

There are other ways to determine associations between verbs and constructions. In an analysis of English tense–aspect (TA) constructions—specifically, of constructions with main verbs marked for the progressive (*–ing*) and past tense (*–ed*)—Wulff et al. (this issue) advocate eschewing simple conditional probabilities in favor of MDCA (Stefanowitsch & Gries, 2003). MDCA calculates a verb's degree of bias toward (or away from) a given construction as compared to a fixed set of alternative constructions.

The two measures—conditional probabilities and MDCA—can suggest two different sets of verbs as highly significant for constructions. The relevant verbs for the progressive construction, for instance, are provided in Table 1.

Wulff et al.'s (this issue) move to consider the MDCA measure is made at least in part because it turns out that many of the top 10 most frequent verbs in the tense–aspect constructions they

TABLE 1
Verbs Most Strongly Associated With the Progressive Construction According to Conditional Probability and MDCA in Two Corpora

	Conditional Probability		MDCA	
Rank	Verb	Corpus	Verb	Corpus
1	Go	BNC & MICASE	Look	BNC & MICASE
2	Do	BNC & MICASE	Sit	BNC & MICASE
3	Be	BNC & MICASE	Play	BNC & MICASE
4	Say	BNC & MICASE	Come	BNC
			Try	MICASE
5	Get	BNC & MICASE	Accord	BNC
			Work	MICASE
6	Come	BNC & MICASE	Wait	BNC
			Go	MICASE
7	Have	BNC & MICASE	Walk	BNC
			Move	MICASE
8	Talk	BNC & MICASE	Joke	BNC
			Sit	MICASE
9	Try	BNC & MICASE	Run	BNC
			Wonder	MICASE
10	Use	MICASE	Watch	BNC
	Look	BNC	Deal	MICASE

Note. All BNC references are to the spoken section of the BNC. BNC = British National Corpus; MDCA = multiple distinctive collexeme analysis; MICASE = Michigan Corpus of Academic Spoken English.

consider are overlapping. For example, *go*, *do*, *be*, *say*, *get*, and *have* are among the lexemes that are most frequent in both the progressive and the past tense constructions (see Wulff et al., Table 3). It is thus not possible to simply say that the verbs that occur most frequently in a construction are the ones that are most strongly associated with it. Instead, the MDCA measure suggests that it is *distinctive* associations that matter: If a verb is strongly associated with one construction but not others, it then potentially facilitates acquisition of that construction. For the TA constructions that Wulff et al. consider, the argument is that verbs that are lexically biased in favor of the progressive (as determined by MDCA) play a special role in its acquisition because they are more homogeneous: As a group, their lexical aktionsart is judged to be relatively atelic.

Lexical bias, as determined by the MDCA measure or by other means, is often highly dependent on the chosen corpus (Hare, McRae, & Elman, 2004). In fact, as is clear in Table 1 for the progressive, only 3 verbs (*looking, sitting, playing*) are among the top 10 maximally biased verbs in both corpora. For the past tense, only 1 verb (*become*) is among the top 10 maximally biased verbs (according to MDCA) in the two corpora. The average overlap across the four TA categories is 33% (comparing Wulff et al., this issue, Tables 5 and 6). There is a greater degree of overlap between the two corpora when conditional probabilities are considered, with 9 out of 10 verbs being most highly frequent in the progressive construction (see Table 1); the average overlap across the four categories is 85% (Wulff et al., Tables 3 and 4).

It is clear that measures of lexical bias are relevant to online sentence processing, as has been shown in numerous studies (Garnsey, Pearlmutter, Myers, & Lotocky, 1997; MacDonald, Pearlmutter, & Seidenberg, 1994). However, the data in this issue do not directly address the question of whether conditional probabilities or lexical bias (or both) are relevant to construction learning. Wulff et al. (this issue) demonstrate that the verbs that display a strong bias toward progressive marking over perfective marking tend to be verbs that are, in neutral contexts and in infinitival forms, judged to be more atelic, and this, of course, makes sense. It is not clear, however, that the same verbs play a crucial role in the formation of the progressive construction.[1]

It could be that the meaning of –*ing* is more transparently available from generalizing the most available (frequent) instances of the construction (the verbs in the far left-hand column of Table 1). It is conceivable that the fact that many of these verb roots alternatively appear with other tense–aspect morphology might even facilitate learning insofar as the learner could use the contrasts in form and meaning to assign function to the morphology. For example, *was doing* differs from *did* in that it is atelic, and *is doing* differs from *do* in that it is atelic. These contrasts could in principle help learners attribute the atelic function to the –*ing* morpheme.

A third possibly relevant measure of words' associations with particular constructions comes from introspective judgments (Ellis & Ferreira-Junior, this issue; Wulff et al., this issue). Ellis and Ferreira-Junior, for example, asked native speakers of English to rate the strength of association that *go*, *put*, *give*, and other verbs have with the meanings typically ascribed to the intransitive motion, the caused-motion, and the ditransitive constructions, respectively. They found that whereas *go*, *put*, and *give* were very strongly associated with their respective constructional meanings, they were not necessarily the verbs that were rated to be the most highly associated with each meaning. For example, participants rated *walk* as being numerically (although not significantly) better than *go* at describing "the movement of something or someone (X) to a new place (Y) or in a new direction"; they likewise rated *push* numerically higher than *put* in describing the meaning, "someone (A) causes the movement of something (X) to a new place (Y) or in a new direction" (Ellis & Ferreira-Junior, Figure 8). *Give* was given the highest rating in describing "how someone (A) causes someone (B) to receive something (Z)," but it was tied with *send*, which occurs in the construction less frequently. In the case of *go* versus *walk*, this might be due to the fact that *walk* is more uniquely associated with the intransitive motion construction than *go*; *go* is used at least as frequently as a future marker. Thus, if both the future and the intransitive construction were taken into account, the MDCA could shed some light on the introspective rankings. The same explanation is not available for *put* as compared with *push*, however, because it is *put* that is strongly biased toward the caused-motion construction; *push* readily appears in other constructions:

(1) a. She put it on the table.
    b. *She put it.
    c. *She put on the table.

(2) a. She pushed it onto the table.
    b. She pushed it.
    c. She pushed on the table.

Thus, the numerical ratings gathered by Ellis and Ferreira-Junior are not systematically correlated with either conditional probabilities or by the MDCA measure. Further research is needed to tease apart the role of each of these possibly relevant factors.

DOES SKEWED INPUT FACILITATE L2 CONSTRUCTION LEARNING?

Thus far we have seen evidence in the present issue that high type frequency has a facilitative effect on L2 acquisition (Collins et al., this issue; McDonough & Kim, this issue). Likewise, skewed input has been argued to be important, especially in the formation of constructional prototypes (Ellis & Ferreira-Junior, this issue; Wulff et al., this issue). Although the studies by Ellis and Ferreira-Junior and Wulff et al. that are discussed in the preceding section provide compelling evidence that skewed constructional input is widely available (as determined by conditional probabilities and also by the MDCA measure), they do not show that skewed input actually *facilitates* construction learning. Establishing this point requires a design that manipulates the amount of variability present in the input and then compares participants' performance on balanced input versus skewed input, as has been done in some modeling and experimental work (Borovsky & Elman, 2006; Casenhiser & Goldberg, 2005; Goldberg et al., 2004). Year and Gordon (this issue) and Year (2009) describe just such a study, and the data that they report indicate a potentially complex relationship between learners' input and their performance on linguistic tasks.

The Year and Gordon (this issue) method involved providing two groups of Korean middle school students with exposure to English dative constructions (i.e., ditransitives and prepositional datives). The groups differed according to the amount of variability present in their input samples. As in Goldberg et al. (2004), overall type and token frequency were held constant, but exposure in the skewed group consisted of input in which one verb (here, *give*) occurred much more frequently than the others. Exposure in the balanced group, however, consisted of input in which all verbs occurred the same number of times. After exposure, the two groups' knowledge of the ditransitive construction was tested on a production task and an acceptability judgment task.

The Year and Gordon (this issue) design took advantage of the fact that ditransitive use is subject to a number of semantic constraints. These are discussed more fully in their article; we focus here on just one for ease of explication: the *recipient animacy constraint*. The ditransitive construction prototypically refers to events in which an agent transfers possession of a theme argument to an animate recipient. Ditransitive exemplars that do not contain an animate recipient violate this constraint and are thus generally seen as less acceptable:

(3) ??Peter sent Chicago a gift.
(4) Peter sent his boss a gift.

On a reading where *Chicago* refers to an inanimate entity, (3) is marked (there is, of course, a significant improvement if *Chicago* is conceptualized as animate; e.g., as the collection of individuals who work in a corporation's Chicago office). In contrast, when the recipient is realized by a referent that is unambiguously animate—for example, *his boss* in (4)—there is no question about acceptability: Sentence (4) is perfectly felicitous.

Knowledge of the recipient animacy constraint should be reflected in participants' behavior. If their exposure to exemplars of the ditransitive has provided them with native-like competence with regard to recipient animacy, then they should, when asked to describe a situation like the one given in (3), tend to avoid using the marked ditransitive and instead produce a prepositional dative construction (e.g., *Peter sent a gift to Chicago*). Likewise, in an acceptability judgment task, participants should show a preference for sentences that abide by the recipient animacy constraint over those that do not: Sentences like (4) should receive higher ratings than sentences like (3).

Year and Gordon (this issue) discuss production and acceptability judgment data for fully felicitous sentences like (4); that is, they ask, given the opportunity to produce a ditransitive when a ditransitive form is allowed, does the skewed input group produce more ditransitives than the balanced group? Additionally, given the opportunity to rate felicitous ditransitive sentences, does the skewed group assign higher ratings than the balanced group? Better performance by the skewed input group on either measure would be consistent with the results of Goldberg and colleagues (Casenhiser & Goldberg, 2005; Goldberg et al., 2004) and would suggest an advantage for skewed input in construction learning.

What Year and Gordon (this issue) find, however, is that no pairwise comparisons between the skewed and balanced groups show statistically significant differences. Although it is clear that both groups have learned something about the ditransitive construction (they both perform significantly better than controls on the two tasks), neither group appears to have an advantage over the other. On the basis of this series of null results, Year and Gordon conclude that skewed input does not facilitate construction learning.

Such a conclusion must, however, be qualified. Null results are always difficult to make sense of because it is hard to determine, when no difference between groups is found, whether the absence of evidence is evidence of an absence of the distinction. More importantly in the present case, it is difficult to maintain the argument that skewed input does not have any facilitative effect when additional data from the same experiment are considered. Year and Gordon (this issue) only report acceptability judgments from sentences like (4) that *do not* violate semantic constraints on ditransitive use, but Year (2009) reported a wider set of results from the same experiment, including judgments for sentences that *do* violate such constraints, like (5); see also (3):

(5) *Peter moved the library the books.

Clearly, the group that finds items such as (5) *less* acceptable has better internalized the relevant constraints. As 2009 makes clear, at the end of 4 days of exposure, the skewed group performs better in this respect: They assign significantly lower ratings to ungrammatical sentences than the balanced group. The balanced group thus appears to be more prone to overgeneralization: They are significantly less likely to correctly reject ungrammatical sentences. At the same time, after a full 8 weeks of exposure, the advantage for skewed frequency is again null. This suggests that there is a facilitative effect of skewed input in the early stages of construction learning but that, over time, learners exposed to balanced input will ultimately catch up.

Thus, the predominant pattern from the experiment described in Year and Gordon (this issue) and Year (2009) appears to be one in which the balanced-input group is less discriminating than the skewed-input group during the initial stages of acquisition; the skewed-input group appears to be appropriately conservative from early on. However, both groups ultimately approximate the behavior of native speakers of English, as we would expect, because not all constructions display the relevant type of skew.

More research is clearly needed to establish what the boundary conditions are under which skewed input may help. Here we have noted that participants show an early increase in sensitivity to semantic constraints when they are exposed to a skewed-input sample. Although this advantage appears to accrue at early stages of constructional development, skewing may actually hinder development at later stages, when a prototypical constructional exemplar must be elaborated in order to engender productivity. As evidence in favor of this hypothesis, consider again the results of McDonough and Kim (this issue). McDonough and Kim's native speakers of Thai had received a minimum of 7 years of instruction in English and had explicitly covered question formation in the semester prior to being tested. Thus, they had far more experience with English questions than Year and Gordon's participants had with the English ditransitive, and they were at a stage of development in which more variability in the input should have helped. This prediction is, in fact, borne out in McDonough and Kim's results: Participants whose prompts forced them to use many different types of lexical verbs in question formation showed a benefit in their ability to produce well-formed questions at test time. This sort of example acts as an important reminder that it is not the case that skewed, low-variance input is always helpful. Rather, it is undoubtedly true that high-variance input—of the kind that McDonough and Kim created in their high type frequency condition—plays a significant role in acquisition as well.

Although Zipfian distributions are common in language, they are not required for category learning. This is especially apparent when one considers that they do not exist for every construction type. Sethuraman and Goodman (2004), for example, demonstrated that instances of the English transitive construction do not center around a single or a few prototypical instances. Even so, the construction is presumably learned. Furthermore, experimental results from Goldberg et al. (2004) and Casenhiser and Goldberg (2005) demonstrated that whereas participants who were exposed to a skewed-input sample outperform those who were exposed to a balanced sample, both groups still outperform controls. Year and Gordon (this issue) also indicate that participants exposed to balanced and skewed input are equally likely to produce the target construction. Thus, although skewed input is advantageous—perhaps particularly so in tasks that require language users to recognize the semantic constraints on a construction—it is not required for learning to occur.

CONCLUSION

We have focused our commentary for the most part on two properties of the input that have been shown to facilitate the acquisition of phrasal constructions: type frequency and skewed input.

Increased type frequencies have been correlated with ease of acquisition (Collins et al., this issue) and with the *elaboration* of item-specific constructional schemas into abstract patterns (Boyd, 2007; Bybee, 1985, 1995; McDonough & Kim, this issue; Plunkett & Marchman, 1991, 1993). Skewed input, however, has been implicated in the initial *creation* of constructional categories (Ellis & Ferreira-Junior, this issue; Wulff et al., this issue)—utterances that bear surface similarities are more likely to be grouped together (Gentner & Medina, 1998)—and semantic constraints are more readily taken into account (Year, 2009). In some respects, these two properties of the input appear to be opposed to one another, as increased type frequency generally goes hand in hand with increased variability.

This is not necessarily true, however. Verbs that are distributed in Zipfian fashion within a construction allow for high type frequency and low category variability to exist side by side in the same input sample (Ellis & Ferreira-Junior, this issue; Goldberg et al., 2004; Wulff et al., this issue). Further, recent experimental evidence indicates that even though Zipfian input puts much of the focus on a handful of high-frequency verbs, learners are still exposed to enough different verb types to drive the development of the sorts of abstract representations necessary for full productivity: Exposure to low-variance, Zipfian input results in the ability to comprehend novel exemplars of newly learned constructions and to use these constructions to describe novel events (Boyd, Gottschalk, & Goldberg, 2009).

It may seem fortuitous that learners' constructional input often seems to have exactly the right distribution of verbs to maximally facilitate acquisition. This aspect of the input is unsurprising, however, when one considers that Zipfian input distributions appear to be an emergent characteristic of lower level semantic and pragmatic verb properties. Speakers use a relatively small number of general-purpose verbs very frequently. When these same verbs have meanings that are coincident with those of a construction—in the way that *go*, *put*, and *give* have meanings that are largely synonymous with the intransitive motion, caused-motion, and ditransitive constructions, for example—Zipfian distributions are a natural outcome.

Finally, all of the articles in the present issue are emblematic of constructionist theory's commitment to developing a model of language acquisition that relies wherever possible on domain-general mechanisms to explain how learning proceeds. We have seen here that language learning is heavily dependent on the structure of the input. Learning is facilitated when, for example, grammatical cues are perceptually salient (Collins et al., this issue) and when the input has an optimal frequency distribution (Collins et al.; Ellis & Ferreira-Junior, this issue; McDonough & Kim, this issue; Wulff et al., this issue; Year, 2009). Note that neither of these properties implicates a learning theory that requires innate, specifically, linguistic principles in order to operate. Instead, learners are able to rely on standard perceptual and categorization abilities that are prevalent throughout human cognition.

ACKNOWLEDGMENTS

We thank Steffi Wulff, JungEun Year, and an anonymous reviewer for helpful discussion and comments.

NOTE

[1]The first words used by children in the progressive construction, as discussed by Andersen and Shirai (1994), include only one verb (*going*) from Table 1. Whether a verb is among those that are the first to be uttered by children (or recorded by researchers) is determined by a number of factors (see Goldberg, 2006; Tomasello & Stahl, 2004).

REFERENCES

Andersen, R. W., & Shirai, Y. (1994). Discourse motivations for some cognitive acquisition principles. *Studies in Second Language Acquisition*, *16*, 133–156.

Bencini, G., & Valian, V. V. (2008). Abstract linguistic representations in 3-year-olds: Evidence from language production and comprehension. *Journal of Memory and Language*, *59*, 97–113.

Borovsky, A., & Elman, J. (2006). Language input and semantic categories: A relation between cognition and early word learning. *Journal of Child Language*, *33*, 759–790.

Boyd, J. K. (2007). Comparatively speaking: A psycholinguistic study of optionality in grammar. Unpublished doctoral dissertation. University of California, San Diego, La Jolla.

Boyd, J. K., Gottschalk, E. A., & Goldberg, A. E. (2009). Linking rule acquisition in novel phrasal constructions. *Language Learning, 59* (Suppl. 1).

Bybee, J. L. (1985). *Morphology: A study of the relation between meaning and form*. Amsterdam: Benjamins.

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

Casasola, M. (2005). When less is more: How infants learn to form an abstract categorical representation of support. *Child Development*, *76*, 279–290.

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

Chomsky, N. (1995). *The minimalist program*. Cambridge, MA: MIT Press.

Christiansen, M., & Chater, N. (2008). Language as shaped by the brain. *Behavioral and Brain Sciences*, *31*, 489–509.

Clark, E. V. (1996). Early verbs, event types and inflections. In C. E. Johnson & J. H. V. Gilbert (Eds.), *Children's language* (pp. 61–73). Mahwah, NJ: Erlbaum.

Collins, L., Trofimovich, P., White, J., Cardoso, W., & Horst, M. (2009). Some input on the easy/difficult grammar question: An empirical study. *Modern Language Journal, 93,* 336–353.

Culicover, P. W., & Jackendoff, R. (2005). *Simpler syntax*. Oxford: Oxford University Press.

Deacon, T. (1997). *The symbolic species: The co-evolution of language and the brain*. New York: W. W. Norton.

Elio, R., & Anderson, J. R. (1984). The effects of information order and learning mode on schema abstraction. *Memory and Cognition*, *12*, 20–30.

Ellis, N. C., & Ferreira-Junior, F. (2009). Construction learning as a function of frequency, frequency distribution, and function. *Modern Language Journal, 93,* 370–385.

Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA: MIT Press.

Fillmore, C. J., Kay, P., & O'Connor, M. C. (1988). Regularity and idiomaticity in grammatical constructions: The case of *let alone*. *Language*, *64*, 501–538.

Garnsey, S. M., Pearlmutter, N. J., Myers, E., & Lotocky, M. (1997). The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *Journal of Memory and Language*, *37*, 58–93.

Gentner, D., Loewenstein, J., & Hung, B. T. (2007). Comparison facilitates children's learning of names for parts. *Journal of Cognition and Development*, *8*, 285–307.

Gentner, D., & Medina, J. (1998). Similarity and the development of rules. *Cognition*, *65*, 263–297.

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

Goldberg, A. E. (1996). Making one's way through the data. In M. Shibatani & S. Thompson (Eds.), *Grammatical constructions: Their form and meaning* (pp. 29–53). Oxford, UK: Clarendon Press.

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

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

Goldberg, A., Casenhiser, D., & White, T. (2007). Constructions as categories of language. *New Ideas in Psychology*, *25*, 70–86.

Hare, M., McRae, K., & Elman, J. (2004). Admitting that admitting verb sense into corpus analyses makes sense. *Language and Cognitive Processes*, *19*, 181–224.

Huttenlocher, J., Vasilyeva, M., & Shimpi, P. (2004). Syntactic priming in young children. *Journal of Memory and Language*, *50*, 182–195.

Kay, P., & Fillmore, C. (1999). Grammatical constructions and linguistic generalizations: The *What's X Doing Y*? construction. *Language*, *75*, 1–33.

Kidd, E., Lieven, E., & Tomasello, M. (2006). Examining the role of lexical frequency in the acquisition and processing of sentential complements. *Cognitive Development*, *21*, 93–107.

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

Lewis, J. D., & Elman, J. L. (2001). A connectionist investigation of linguistic arguments from the Poverty of the Stimulus: Learning the unlearnable. In *Proceedings of the 23rd annual conference of the Cognitive Science Society* (pp. 552–557). Mahwah, NJ: Erlbaum.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, *101*, 676–703.

Maguire, M. J., Hirsh-Pasek, K., Golinkoff, R. M., & Brandone, A. C. (2008). Focusing on the relation: Fewer exemplars facilitate children's initial verb learning and extension. *Developmental Science*, *11*, 628–634.

McDonough, K., & Kim, Y. (2009). Syntactic priming, type frequency, and EFL learners' production of *wh*-questions. *Modern Language Journal*, *93,* 386–398.

Mervis, C., & Rosch, E. (1981). Categorization of natural objects. *Annual Reviews in Psychology*, *32*, 89–115.

Mintz, T. H., Newport, E. L., & Bever, T. (2002). The distributional structure of grammatical categories in speech to young children. *Cognitive Science*, *26*, 393–424.

Ninio, A. (1999). Pathbreaking verbs in syntactic development and the question of prototypical transitivity. *Journal of Child Language*, *26*, 619–653.

Ninio, A. (2005). Accelerated learning without semantic similarity: Indirect objects. *Cognitive Linguistics*, *16*, 531–556.

Onnis, L., Waterfall, H. R., & Edelman, S. (2008). Learn locally, act globally: Learning language from variation set cues. *Cognition*, *109*, 423–430.

Perfors, A., Tenenbaum, J. B., & Regier, T. (2006). Poverty of the stimulus? A rational approach. In *Proceedings of the 28th annual conference of the Cognitive Science Society* (pp. 663–668). Mahwah, NJ: Erlbaum.

Plunkett, K., & Marchman, V. (1991). U-shaped learning and frequency effects in a multi-layered perceptron: Implications for child language acquisition. *Cognition*, *38*, 43–102.

Plunkett, K., & Marchman, V. (1993). From rote learning to system building: Acquiring verb morphology in children and connectionist nets. *Cognition*, *48*, 21–69.

Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology*, *77*, 353–363.

Ramscar, M., & Yarlett, D. (2007). Linguistic self-correction in the absence of feedback: A new approach to the logical problem of language acquisition. *Cognitive Science*, *31*, 927–960.

Savage, C., Lieven, E., Theakston, A., & Tomasello, M. (2003). Testing the abstractness of children's linguistic representations: Lexical and structural priming of syntactic constructions in young children. *Developmental Science*, *6*, 557–567.

Savage, C., Lieven, E., Theakston, A., & Tomasello, M. (2006). Structural priming as implicit learning in language acquisition: The persistence of lexical and structural priming in 4-year-olds. *Language Learning and Development*, *2*, 27–49.

Sethuraman, N. (2002). *The acquisition of verbs and argument structure constructions*. Unpublished doctoral dissertation, University of California, San Diego, La Jolla.

Sethuraman, N., & Goodman, J. (2004). Children's mastery of the transitive construction. In E. Clark (Ed.), *Online proceedings of the 32nd session of the Stanford Child Language Research Forum* (pp. 60–67). Stanford, CA: CSLI Publications.

Slobin, D. I. (1982). Universal and particular in the acquisition of language. In E. Wanner & L. R. Gleitman (Eds.), *Language acquisition: The state of the art* (pp. 129–170). New York: Cambridge University Press.

Stefanowitsch, A., & Gries, S. T. (2003). Collostructions: Investigating the interaction between words and constructions. *International Journal of Corpus Linguistics*, *8*, 209–243.

Thompson, S. A. (2002). "Object complements" and conversation: Towards a realistic account. *Studies in Language*, *26*, 125–163.

Tomasello, M. (2000). The item-based nature of children's early syntactic development. *Trends in Cognitive Sciences*, *4*, 156–163.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

Tomasello, M., & Stahl, D. (2004). Sampling children's spontaneous speech: How much is enough? *Journal of Child Language*, *31*, 101–121.

Wulff, S., Ellis, N. C., Römer, U., Bardovi-Harlig, K., & LeBlanc, C. (2009). The acquisition of tense–aspect: Converging evidence from corpora, cognition, and learner constructions. *Modern Language Journal*, *93,* 354–369.

Year, J. (2009). *Korean speakers' acquisition of the English ditransitive construction: The role of input frequency and distribution*. Unpublished doctoral dissertation. Columbia University, New York.

Year, J., & Gordon, P. (2009). Korean speakers' acquisition of the English ditransitive construction: The role of verb prototype, input distribution, and frequency. *Modern Language Journal*, *93,* 399–417.

Zipf, G. K. (1935). *The psycho-biology of language*. New York: Houghton-Mifflin.