



Subtle Implicit Language Facts Emerge from the Functions of Constructions

Adele E. Goldberg*

Department of Psychology, Princeton University, Princeton, NJ, USA

Much has been written about the unlikelihood of innate, syntax-specific, universal knowledge of language (Universal Grammar) on the grounds that it is biologically implausible, unresponsive to cross-linguistic facts, theoretically inelegant, and implausible and unnecessary from the perspective of language acquisition. While relevant, much of this discussion fails to address the sorts of facts that generative linguists often take as evidence in favor of the Universal Grammar Hypothesis: subtle, intricate, knowledge about language that speakers implicitly know without being taught. This paper revisits a few often-cited such cases and argues that, although the facts are sometimes even more complex and subtle than is generally appreciated, appeals to Universal Grammar fail to explain the phenomena. Instead, such facts are strongly motivated by the *functions of the constructions* involved. The following specific cases are discussed: (a) the distribution and interpretation of anaphoric *one*, (b) constraints on long-distance dependencies, (c) subject-auxiliary inversion, and (d) cross-linguistic linking generalizations between semantics and syntax.

Keywords: anaphoric one, island constraints, subject-auxiliary inversion, universal grammar, grammatical constructions

INTRODUCTION

We all recognize that humans have a different biological endowment than the prairie vole, the panther, and the grizzly bear. We can also agree that only humans have human-like language. Finally, we agree that adults have representations that are specific to language (for example, their representations of constructions). The question that the present volume focuses on is whether we need to appeal to representations concerning syntax that have not been learned in the usual way—that is on the basis of external input and domain-general processes—in order to account for the richness and complexity that is evident in all languages. The Universal Grammar Hypothesis is essentially a claim that we do. It asserts that certain syntactic representations are “innate,”¹ in the sense of not being learned, and that these representations both facilitate language acquisition and constrain the structure of all real and possible human languages².

I take this Universal Grammar Hypothesis to be an important empirical claim, as it is often taken for granted by linguists and it has captured the public imagination. In particular, linguists

¹I put the term “innate” in quotes because the term lacks an appreciation of the typically complex interactions between genes and the environment before and after birth (see Deák, 2000; Blumberg, 2006; Karmiloff-Smith, 2006 for relevant discussion).

²Universal Grammar seems to mean different things to different researchers. In order for it to be consistent with its nomenclature and its history in the field, I take the Universal Grammar Hypothesis to claim that there exists some sort of universal but unlearned (“innate”) knowledge of language that is specific to grammar.

OPEN ACCESS

Edited by:

N. J. Enfield,
University of Sydney, Australia

Reviewed by:

Jeffrey Lidz,
University of Maryland, USA
Elizabeth Closs Traugott,
Stanford University, USA

*Correspondence:

Adele E. Goldberg
adele@princeton.edu

Specialty section:

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

Received: 26 October 2015

Accepted: 17 December 2015

Published: xx January 2016

Citation:

Goldberg AE (2016) Subtle Implicit
Language Facts Emerge from the
Functions of Constructions.
Front. Psychol. 6:2019.
doi: 10.3389/fpsyg.2015.02019

often assume that infants bring with them to the task of learning language, knowledge of noun, verb, and adjective categories, a restriction that all constituents must be binary branching, a multitude of inaudible but meaningful “functional” categories and placeholders, and constraints on possible word orders. This is what Pearl and Sprouse seem to have in mind when they note that positing Universal Grammar to account for our ability to learn language is “theoretically unappealing” in that it requires learning biases that “appear to be an order (or orders) of magnitude more complex than learning biases in any other domain of cognition” (Pearl and Sprouse, 2013, p. 24).

The present paper focuses on several phenomena that have featured prominently in the mainstream generative grammar literature, as each has been assumed to involve a purely syntactic constraint with no corresponding functional basis. When constraints are viewed as arbitrary in this way, they appear to be mysterious and are often viewed as posing a learnability challenge; in fact, each of the cases below has been used to argue that an “innate” Universal Grammar is required to provide the constraints to children *a priori*.

The discussion below aims to demystify the restrictions that speakers implicitly obey, by providing explanations of each constraint in terms of the functions of the constructions involved. That is, constructions are used in certain constrained ways and are combined with other constructions in constrained ways, because of their semantic and/or discourse functions. Since children *must* learn the functions of each construction in order to use their language appropriately, the constraints can then be understood as emerging as by-products of learning those functions. In each case, a generalization based on the communicative functions of the constructions is outlined and argued to capture the relevant facts better than a rigid and arbitrary syntactic stipulation (see also DuBois, 1987; Hopper, 1987; Michaelis and Lambrecht, 1996; Kirby, 2000; Givón, 2001; Auer and Pfänder, 2011). Thus, recognizing the functional underpinnings of grammatical phenomena allows us to account for a wider, richer range of data, and allows for an *explanation* of that data in a way that purely syntactic analyses do not.

In the following sections, functional underpinnings of the distribution and interpretation of various constructions are offered including anaphoric *one*, various long-distance dependences, subject-auxiliary inversion, and cross-linguistic linking generalizations.

ANAPHORIC ONE

Anaphoric *One*'s Interpretation³

There are many interesting facts of language; let's consider this *one*. The last word in the previous sentence refers to an “interesting fact about language” in the first clause; it cannot refer to an interesting fact that is about something other than language. This type of observation has been taken to imply that *one* anaphora demonstrates “innate” knowledge that full noun

³This section is based on Goldberg and Michaelis (2015), which contains a much more complete discussion of anaphoric *one* and its relationship to numeral *one* (and other numerals).

phrases (or “DP”s) contain a constituent that is larger than a noun but smaller than a full noun phrase: an N' (*interesting fact of language* above), and, that *one* anaphora must refer to an N', and may never refer to a noun without its grammatical complement (Baker, 1978; Hornstein and Lightfoot, 1981; Radford, 1988; Lidz et al., 2003b). However, as many researchers have made clear, anaphoric *one* actually *can* refer to a noun without its complement as it does in the following attested examples from the COCA corpus (Davies, 2008; for additional examples and discussion see Lakoff, 1970; Jackendoff, 1977; Dale, 2003; Culicover and Jackendoff, 2005; Payne et al., 2013; Goldberg and Michaelis, 2015)⁴.

1. “not only would **the problem of alcoholism** be addressed, but also **the related one of violence**,” [smallest N' = *problem of alcoholism*; but *one* = “problem”]
2. “it was a **war of choice** in many ways, not **one of necessity**.” [smallest N' = *war of choice*; *one* = “war”]
3. “Turning a **sense of ostracism** into **one of inclusion** is a difficult trick. [smallest N' = *sense of ostracism*; *one* = “sense”]
4. “more a **sign of desperation** than **one of strength**” [smallest N' = *sign of desperation*; *one* = “sign”]

In each case, the “*of* phrase” (e.g., *of alcoholism* in 1) is a complement according to standard assumptions and therefore should be included in the smallest available N' that the syntactic proposal predicts *one* can refer to. Yet in each case, *one* actually refers only to the previous noun (*problem*, *war*, *sense*, and *sign*, respectively, in 1–4), and does not include the complement of the noun.

In the following section, I outline an explanation of *one*'s distribution and interpretation, which follows from its discourse function. To do this, it is important to appreciate anaphoric *one*'s close relationship to numeral *one*, as described below.

The Syntactic and Semantic Behavior of *One* are Motivated by its Function

Leaving aside the wide range of linguistic and non-linguistic entities that *one* can refer to for a moment, let us consider the linguistic contexts in which *one* itself occurs. Goldberg and Michaelis (2015) observe that anaphoric *one* has the same grammatical distribution as numeral *one* (and other numerals), when the latter are used without a head noun. The only formal distinction between anaphoric *one* and the elliptical use of numeral *one* is that numeral *one* receives a sentence accent, as indicated by capital letters in **Table 1**, whereas anaphoric *one* must be unstressed (Goldberg and Michaelis, 2015).

The difference in accent between cardinal and anaphoric *one* reflects a key difference in their functions. Whereas cardinal *one* is used to assert the quantity “1,” anaphoric *one* is used when quality or existence—not quantity—is at issue. That is, if asked about quantity as in (5), a felicitous response (5a) involves cardinal *one*, which is necessarily accented (5a; cf. 5b). If the

⁴A version of the first sentence also allows *one* to refer to an interesting *fact* that is not about language:

a. There are many interesting facts of language, but let's consider this *one* about music.

229 *type of entity* is at issue as in (6), then anaphoric *one*, which is
 230 necessarily unaccented, is used (6b; cf. 6a):

- 231 5. Q: *How many dogs does she have?*
 232 a. She has (only) ONE. (cardinal ONE)
 233 b. #She has a big one. (anaphoric one)
 234 6. Q: *What kind of dog does she have*
 235 a. #She has (only) ONE (cardinal ONE)
 236 b. She has a BIG one. (anaphoric one).
 237

238 It is this fact, that anaphoric *one* is used when quality and not
 239 quantity is at issue, that explains why anaphoric *one* so readily
 240 picks out an entity, recoverable in the discourse context, that
 241 often corresponds to an N': anaphoric *one* often refers to a noun
 242 and its complement (or modifier) because the complement or
 243 modifier supplies the quality. But the quality can be expressed
 244 explicitly as it is in (6b; with *big*) or in (1–4) with the overt
 245 complement phrases⁵. If existence (and not quality or quantity) is
 246 at issue, anaphoric *one* can refer to a full noun phrase as in (7):
 247

- 248 7. [Who wants a drink?] I'll take one.

249 Thus, given the fact that anaphoric *one* exists in English,
 250 its semantic relationship to cardinal numeral *one* predicts its
 251 distribution and interpretation. Anaphoric *one* is used when the
 252 quality or existence of an entity evoked in the discourse—not its
 253 cardinality—is relevant.

254 The only additional fact that is required is a representation of
 255 the plural form, *ones*, and both the form and the function of *ones*
 256 is motivated because *ones* is a lexicalized extension of anaphoric
 257 *one* (Goldberg and Michaelis, 2015). *Ones* differs from anaphoric
 258 *one* only in being plural both formally and semantically; like
 259 singular anaphoric *one*, plural *ones* evokes the quality or existence
 260 and not the cardinality of a type of entity recoverable in context.

261 There are several lessons that can be drawn from this simple
 262 case. First, if we are too quick to assume a purely syntactic
 263 generalization without careful attention to attested data, it is
 264 easy to be led astray. Moreover, it is important to recognize
 265 relationships among constructions. In particular, anaphoric *one*
 266 is systematically related to numeral *one*, and a comparison of
 267 the functional properties of these closely related forms serves to
 268 explain their distributional properties.
 269

270
 271 **TABLE 1 | Distributional contexts for anaphoric *one* and the elliptical use**
 272 **of cardinal *one*.**

273 Anaphoric <i>one</i>	274 Numeral <i>one</i> (1)
275 She asked for <u>one</u> .	275 She asked for <u>ONE</u> .
276 She got <u>a blue one</u> .	276 She got <u>a mere ONE</u> .
277 She only wanted <u>that one</u> .	277 She only wanted <u>that ONE</u> .
278 She was <u>one of a group</u> .	278 She was <u>ONE</u> of a group.

280 *The two differ only in that only numeral one receives a sentence accent and asserts the*
 281 *quantity "1."*

282
 283 ⁵To fully investigate the range of data that have been proposed to date in the
 284 literature, judgment data should be collected in which contexts are systematically
 285 varied to emphasize definiteness, quality, existence and cardinality.

286 There remain interesting questions about how children learn
 287 the function of anaphoric *one*. But once we acknowledge that
 288 children *do* learn its function—and they must in order to use it
 289 in appropriate discourse contexts—there is nothing mysterious
 290 about its formal distribution.
 291

292 CONRAINTS ON LONG DISTANCE 293 DEPENDENCIES 294

295 The Basic Facts 296

297 Most languages allow constituents to appear in positions other
 298 than their most canonical ones, and sometimes the distance
 299 between a constituents' actual position and its canonical position
 300 can be quite long. For example, when questioned, the phrase
 301 *which/that coffee* in (8) is not where it would appear in a canonical
 302 statement; instead, it is positioned at the front of the sentence,
 303 and there is a *gap* (indicated by "____") where it would normally
 304 appear.
 305

- 306 8. Which coffee did Pam say Sam likes ____ better than tea?
 307 (cf. Pam said Sam likes that coffee better than tea.)
 308

309 This type of relationship is often discussed as if the constituent
 310 "moved" or was "extracted" from its canonical position, although
 311 no one has believed since Fodor et al. (1974) that the movement is
 312 anything more than a metaphor. I use more neutral terminology
 313 here and refer to the relation between the actual position and the
 314 canonical position as a long-distance dependency (LDD).
 315

316 There are several types of LDD constructions including wh-
 317 questions, the topicalization construction, cleft constructions,
 318 and relative clause constructions. These are exemplified in
 319 **Table 2**.

320 Ross (1967) long ago observed that certain other types of
 321 constructions resist containing the gap of a LDD. That is,
 322 certain constructions are "islands" from which constituents
 323 cannot escape. Combinations of an "island construction"
 324 with a LDD construction result in ill-formedness (see
 325 **Table 3**):
 326

327
 328 **TABLE 2 | Examples of long distance dependency (LDD) constructions:**
 329 **constructions in which a constituent appears in a fronted position instead**
 330 **of where it would canonically appear.⁶**

331 Wh-questions	331 <u>What</u> did Pam say Sam likes ____ better than tea?
332 Topicalization construction	332 <u>That coffee</u> , Pam said Sam likes ____ better than tea.
333 It-cleft construction	333 It was <u>that coffee</u> that Pam said Sam likes ____ better than tea.
334 Relative Clause construction	334 She tasted <u>the coffee</u> that Pat said Sam likes ____ better than tea.

335
 336
 337
 338 ⁶Other LDD constructions include comparatives (Bresnan, 1972; Merchant, 2009)
 339 and "tough" movement constructions (Postal and Ross, 1971) which should fall
 340 under the present account as well; more study is needed to investigate these cases
 341 systematically from the current perspective (see Hicks (2003); Sag (2010); for
 342 discussion).

TABLE 3 | Examples of island constructions: constructions that resist containing the gap in a LDD (Ross, 1967).

343	??Who did [that she hit _] was horrible?	Subjects
344	(cf. [That she hit him] was horrible.)	
345	??Who did she see [the boy who met _]?	
346	(cf. She saw [the boy who met Sally])	(Most) complex noun phrases
347		
348	??What did she read [the letter that was about_]?	
349	(cf. She read [the letter that was about mountains])	
350	??Who did she mumble/deny [he liked_]?	Clausal complements of manner of speaking and factive verbs
351	(cf. She mumbled/denied [he liked Katherine]).	
352		
353	??Who did she eat spaghetti [while talking to _on the phone]?	(Certain) adjuncts
354	(cf. She ate spaghetti [while talking to Jenny on the phone]).	
355		
356		
357		

A Clash Between the Functions of LDD Constructions and the Functions of Island Constructions

Several researchers have observed that INFORMATION STRUCTURE plays a key role in island constraints (Takami, 1989; Deane, 1991; Engdahl, 1997; Erteschik-Shir, 1998; Polinsky, 1998; Van Valin, 1998; Goldberg, 2006, 2013; Ambridge and Goldberg, 2008). Information structure refers to the way that information is “packaged” for the listener: constituents are topical in the discourse, part of the potential focus domain, or are backgrounded or presupposed (Halliday, 1967; Lambrecht, 1994). Different constructions that convey “the same thing,” typically exist in a given language in order to provide different ways of packaging the information, and thus information structure is perhaps the most important reason why languages have alternative ways to say the “same” thing. As explained below, the ill-formedness of island effects arises essentially from a clash between the function of the LDD construction and the function of the island construction. First, a few definitions are required.

The FOCUS DOMAIN is that part of a sentence that is asserted. It is thus “one kind of emphasis, that whereby the speaker marks out a part (which may be the whole) of a message block as that which he wishes to be interpreted as informative” Halliday (1967: 204). Similarly Lambrecht (1994: 218) defines the focus relation as relating “the pragmatically non-recoverable to the recoverable component of a proposition [thereby creating] a new state of information in the mind of the addressee.” What parts of a sentence fall within the focus domain can be determined by a simple negation test: when the main verb is negated, only those aspects of a sentence within the potential focus domain are negated. Topics, presupposed constituents, constituents within complex noun phrases, and parenthetical remarks are not part of the focus domain, as they are not negated by sentential negation.⁷

⁷Backgrounded constituents can be negated with “metalinguistic” negation, signaled by heavy lexical stress on the negated constituent (*I didn't read the book that Maya gave me because she didn't GIVE me any book!*). But then metalinguistic negation can negate anything at all, including intonation, lexical choice, or accent. Modulo this possibility, the backgrounded constituents of a sentence are not part of what is asserted by the sentence.

9. Pam, as I told you before, didn't sell the book to the man she just met.

→ negates that the book was sold; does not negate that she just met a man or that the speaker is repeating herself.

It has long been observed that the gap in a LDD construction is typically within the potential focus domain of the utterance (Takami, 1989; Erteschik-Shir, 1998; Polinsky, 1998; Van Valin, 1998; see also Morgan, 1975): this predicts that topics, presupposed constituents, constituents within complex noun phrases, and parentheticals are all island constructions and they are (see previous work and Goldberg, 2013 for examples).

It is necessary to expand this view slightly by defining BACKGROUNDED CONSTITUENTS to include everything in a clause except constituents within the focus domain *and the subject*. Like the focus domain, the subject argument is part of what is made prominent or foregrounded by the sentence in the given discourse context, since the subject argument is the default TOPIC of the clause or what the clause is “about” (MacWhinney, 1977; Chafe, 1987; Langacker, 1987; Lambrecht, 1994). That is, a clausal topic is a “matter of [already established] current interest which a statement is about and with respect to which a proposition is to be interpreted as relevant” (Michaelis and Francis, 2007: 119). The topic serves to contextualize other elements in the clause (Strawson, 1964; Kuno, 1976; Langacker, 1987; Chafe, 1994). We can now state the restriction on LDDs succinctly:

★ Backgrounded constituents cannot be “extracted” in LDD constructions (Backgrounded Constituents are Islands; Goldberg, 2006, 2013).

The claim in ★ entails that only elements within the potential focus domain *or* the subject are candidates for LDDs. Notice that constituents properly contained *within* the subject argument are backgrounded in that they are not themselves the primary topic, nor are they part of the focus domain. Therefore, subjects are “islands” to extraction.

Why should ★ hold? The restriction follows from a clash of the functions of LDD constructions and island constructions. As explained below: a referent cannot felicitously be both discourse-prominent (in the LDD construction) and backgrounded in discourse (in the island construction). That is, LDD constructions exist in order to position a particular constituent in a discourse-prominent slot; island constructions ensure that the information that they convey is backgrounded in discourse. It is anomalous for an argument, which the speaker has chosen to make prominent by using a LDD construction, to correspond to a gap that is within a backgrounded (island) construction.

What is meant by a discourse-prominent position? The *wh*-word in a question LDD is a classic focus, as are the fronted elements in “cleft” constructions, another type of LDD. The fronted argument in a topicalization construction is a newly established topic (Gregory and Michaelis, 2001)⁸. Each of these LDD constructions operates at the sentence level and the main

⁸The present understanding of discourse prominence implicitly acknowledges the notions of topic and focus are not opposites: both allow for constituents to be interpreted as being prominent (see, e.g., Arnold, 1998: for experimental and corpus evidence demonstrating the close relationship between topic and focus).

457 clause topic and focus are classic cases of discourse-prominent
458 positions.

459 The relative clause construction is a bit trickier because the
460 head noun of a relative clause—the “moved” constituent—is
461 not necessarily the main clause topic or focus, and so it may
462 not be prominent in the general discourse. For this reason, it
463 has been argued that relative clauses involve a case of recycling
464 the formal structure and constraints that are motivated in
465 the case of questions to apply to a distinct but related case:
466 relative clauses (Polinsky, 1998). But in fact, the head noun in
467 a relative clause construction is prominent when it is considered
468 *in relation to* the relative clause itself: the purpose of a relative
469 clause is to identify or characterize the argument expressed
470 by the head noun. In this way, the head noun should not
471 correspond to a constituent that is backgrounded within the
472 relative clause. Thus, there is a clash for the same reason that
473 sentence level LDD constructions clash with island constructions,
474 except that what is prominent and what is backgrounded *is*
475 *relative to* the content of the NP: the head noun is prominent
476 and any island constructions within the relative clause are
477 backgrounded.

478 We should expect the ill-formedness of LDDs to be gradient
479 and degrees of ill-formedness are predicted to correspond
480 to degrees of backgroundedness, when other factors related
481 to frequency, plausibility, and complexity are controlled for.
482 This idea motivated an experimental study of various clausal
483 complements, including “bridge” verbs, manner-of-speaking
484 verbs, and factive verbs and exactly the expected correlation
485 was found (Ambridge and Goldberg, 2008): the degree of
486 acceptability of extraction showed a strikingly strong inverse
487 correlation with the degree of backgroundedness of the
488 complement clause—which was operationalized by judgments
489 on a negation test. Thus, the claim is that each construction
490 has a function and that constructions are combined to form
491 utterances; constraints on “extraction” arise from a clash of
492 discourse constraints on the constructions involved.

493 The functional account predicts that certain cases pattern
494 as they do, even though they are exceptional from a purely
495 syntactic point of view (see also Engdahl, 1997). These include
496 the cases in **Table 4**. Nominal complements of indefinite
497 “picture nouns” fall within the focus domain, as do certain
498 adjuncts, while the recipient argument of the double object
499 construction, as a secondary topic, does not (see Goldberg,
500 2006, 2013 for discussion). Therefore, the first two cases in
501 **Table 2** are predicted to allow LDDs while the final case is
502 predicted to resist LDDs⁹. No special assumptions or stipulations
503 are required.

504 There is much more to say about island effects (see e.g.,
505 Sprouse and Hornstein, 2013). The hundreds of volumes written
506 on the subject cannot be properly addressed in a short review
507 such as this. The goal of this section is to suggest that

508
509 This makes sense once we realize that one sentence’s focus is often the next
510 sentence’s topic.

511 ⁹Cross linguistic work is needed to determine whether secondary topics generally
512 resist LDDs as is the case in the English double-object construction, or whether the
513 dispreference is only detectable when an alternative possibility is available, as in
514 English, where questioning the recipient of the *to*-dative is preferred (see note 10).

TABLE 4 | Cases that follow from an information structure account, but not from an account that attempts to derive the restrictions from configurations of syntactic trees.

Who did she take [a picture of ___]? (cf. She took [a picture of Sally])	Reduced relative clauses that are within the focus domain (e.g., “picture NPs”) are not islands; those that are not within the focus domain are islands.	514 515 516
Who did she wait in line [in order to see ___]? (cf. She waited in line [in order to see U2]).	Non-backgrounded adjuncts	517 518 519 520 521 522 523
??Who did she give the book? ¹⁰ (cf. She gave Aliza the book.) (cf. also, Who did she give the book to?)	Backgrounded (as the secondary topic) recipient argument of the double object construction	524 525 526 527

528
529 a recognition of the functions of the relevant constructions
530 involved can explain which constructions are islands and why;
531 much more work is required to explore whether this proposal
532 accounts for each and every LDD construction in English and
533 other languages.

534 SUBJECT AUXILIARY INVERSION (SAI) 535

536 SAI’s Distribution 537

538 Subject-auxiliary inversion (e.g., *is this it?*) has a distribution that
539 is quite unique to English. In Old English, it followed a more
540 general “verb second” pattern, which still exists in Germanic and
541 a few other languages. But English changed, as languages do, and
542 today, subject-auxiliary inversion requires an *auxiliary* verb and
543 is restricted to a limited range of constructions, enumerated in
544 (10–17):

- | | | |
|--|----------------------------------|------------|
| 10. Did she go? | Y/N questions | 546 |
| Where did she go? | (non-subject) WH-questions | 547 |
| 11. Had she gone, they would be here by now. | Counterfactual conditionals | 548
549 |
| 12. Seldom had she gone there. | Initial negative adverbs | 550 |
| 13. May a million fleas infest his armpits! | Wishes/Curses | 551 |
| 14. He was faster at it than was she. | Comparatives | 552 |
| 15. Neither do they vote. | Negative conjunct | 553 |
| 16. Boy did she go, or what?! | Exclamatives | 554 |
| 17. So does he. | Positive elliptical conjunctions | 555 |

556 When SAI is used, the entire subject argument appears after the
557 first main clause auxiliary as is clear in a comparison of (18a) and
558 (18b):

- 559
560 18. a. Has the girl who was in the back of the room had enough
561 to eat? (inverted).
562 b. The girl who was in the back of the room has had enough
563 to eat. (non-inverted).

564 Notice that the very first auxiliary in the corresponding
565 declarative sentence (*was*) cannot be inverted (see 19a), nor can
566 the second (or other) main clause auxiliary (see 19b).

567
568 ¹⁰Support for this judgment comes from the fact that questions of the recipient of
569 the *to*-dative outnumber those of the recipient of the double-object construction
570 in corpus data by a factor of 40 to 1 (Goldberg, 2006: 136).

- 571 19. a. *Was the girl who in the back of the room has had enough
572 to eat? (only the main clause auxiliary can be inverted).
573 b. *Had the girl who was in the back of the room has enough
574 to eat? (only the *first* main clause auxiliary can be
575 inverted).

576 Thus, the generalization at issue is that the first auxiliary in
577 the full clause containing the subject is inverted with the entire
578 subject constituent.

579 SAI occurs in a range of constructions in English and each one
580 has certain unique constraints and properties (Fillmore, 1999;
581 Goldberg, 2009); for example, in the construction with negative
582 adverbs (e.g., 12), the adverb is positioned clause initially; curses
583 (e.g., 13) are quite particular about which auxiliary may be used
584 (*May a million fleas invest your armpits.* vs. **Might/will/shall a*
585 *million fleas invest your armpits!*); and inversion in comparatives
586 (e.g., 14) is restricted to a formal register. Thus, any descriptively
587 adequate account of SAI in English must make reference to these
588 properties of individual constructions.

589 The English constructions evolved diachronically from a
590 more general constraint which still operative in German main
591 clauses. But differences exist across even these closely related
592 languages. The German constraint applies to main verbs, while
593 English requires an auxiliary verb, and in English the auxiliary is
594 commonly in first not second position (e.g., *did I get that right?*).
595 Also, verb-second in German is a main clause phenomenon,
596 but in English, SAI is possible in embedded clauses as
597 well (20–21):
598

- 599 20. “And Janet, do you think that had he gotten a diagnosis
600 younger, it would have been a different outcome?” (COCA)
601 21. “Many of those with an anti-hunting bias have the idea that
602 were it not for the bloodthirsty human hunter, game would
603 live to ripe old age” (COCA)

604 Simple recurrent connectionist networks can learn to invert
605 the correct auxiliary on the basis of simpler input that
606 children uncontroversially receive (Lewis and Elman, 2001).
607 This model is instructive because it is able to generalize
608 correctly to produce complex questions (e.g., *Is the man*
609 *who was green here?*), after receiving training on simple
610 questions and declarative statements with a relative clause.
611 The network takes advantage of the fact that both simple
612 noun phrases (*the boy*) and complex noun phrases (*The*
613 *boy who chases dogs*) have similar distributions in the input
614 (see also Pullum and Scholz, 2002; Reali and Christiansen,
615 2005¹¹; Ambridge et al., 2006; Rowland, 2007; Perfors et al.,
616 2011).

617 *The reason* simple and complex subjects have similar
618 distributions is that the subject is a coherent semantic unit,
619 typically referring to an entity or set of entities. For example,
620 in (22a–c), *he*, *the boy*, and *the boy in the front row*, all identify
621 a particular person and each sentence asserts that the person in
622 question is tall.
623

624 ¹¹See Kam et al. (2008) for discussion of the difficulties of using only bi-grams.
625 Since we assume that meaningful units are combined to form larger meaningful
626 units, resulting in hierarchical structure, this critique does not undermine the
627 present proposal.

- 22.a. He is tall. 628
b. The boy is tall. 629
c. The boy who sat in front of me is tall. 630

631 Thus the distributional fact that is sufficient for learning the key
632 generalization is that subjects, whether simple or complex, serve
633 the same function in sentences.

634 We might also ask *why* SAI is used in the range of
635 constructions it is, and why these constructions use this formal
636 feature instead of placing the subject in sentence-final position
637 or some other arbitrary feature. Consider the function of the
638 first auxiliary of the clause containing the subject. This auxiliary
639 indicates tense and number agreement (23), but an auxiliary is
640 not required for these functions, as the main verb can equally well
641 express them (24).
642

23. a. She did say. 643
b. They do say. 644

24. a. She said. 645
b. They say. 646

647 The first auxiliary of the clause containing the subject obligatorily
648 serves a different purpose related to negative or emphasized
649 positive polarity (Langacker, 1991). That is, if a sentence is
650 negated, the negative morpheme occurs immediately after—often
651 cliticized to—the first auxiliary of the clause that contains the
652 subject (25):
653

25. She hadn't been there. 654

655 And if positive polarity is emphasized, it is the first auxiliary that
656 is accented (26):
657

26. She HAD been there. (cf. She had been there). 658

659 If the corresponding simple positive sentence does not contain an
660 auxiliary, the auxiliary verb *do* is drafted into service (27):
661

- 27.a. She DID swim in the ocean. 662
b. She did not swim in the ocean. 663
c. She didn't swim in the ocean. 664
(cf. She swam in the ocean). 665

666 Is it a coincidence that the first auxiliary of the main clause
667 that contains the subject conveys polarity? Intriguingly, most
668 SAI constructions offer different ways to implicate a negative
669 proposition, or at least to avoid asserting a simple positive one
670 (Brugman and Lakoff, 1987; Goldberg, 2006)¹². For example,
671 yes/no questions ask whether or not the proposition is true;
672 counterfactual conditionals deny that the antecedent holds; and
673 the inverted clause in a comparative can be paraphrased with a
674 negated clause as in (28):
675

28. He was faster than was she. → She was not as fast as he was. 676

677 Exclamatives have the form of rhetorical yes/no questions, and
678 in fact they commonly contain tag questions (e.g., *Is he a jerk,*
679 *or what?!*) (Goldberg and Giudice, 2005). They also have the
680

681 ¹²Labov (1968) discusses another SAI construction used in AAVE, which requires
682 a negated auxiliary (e.g., *Can't nobody go there.*).
683
684

685 pragmatic force of emphasizing the positive polarity, which we
686 have seen is another function of the first auxiliary. Likewise, the
687 positive conjunction (*so did she*) emphasizes positive polarity as
688 well.

689 Thus the form of SAI in English is motivated by the functions
690 of the vast majority of SAI constructions: in order to indicate
691 non-canonical polarity of a sentence—either negative polarity or
692 emphasized positive polarity—the auxiliary required to convey
693 polarity is inverted. Once the generalization is recognized to be
694 iconic in this way, it becomes much less mysterious both from a
695 descriptive and an acquisition perspective.

696 There is only one case where SAI is used without implicating
697 either negative polarity or emphasizing positive polarity: non-
698 subject *wh*-questions. This case appears to be an instance of
699 recycling a formal pattern for use with a construction that has
700 a related function to one that is directly motivated (see also
701 Nevalainen, 1997). In particular, *wh*-questions have a function
702 that is clearly related to *yes/no* questions since both are questions.
703 But while SAI is directly motivated by the non-positive polarity
704 of *yes/no* questions, this motivation does not extend to *wh*-
705 questions (also see Goldberg, 2006 and Langacker, 2012 for a
706 way to motivate SAI in *wh*-questions more directly). Nonetheless,
707 to ignore the relationship between the function of the first
708 auxiliary as an indicator of negative polarity or emphasized
709 positive polarity, and the functions of SAI constructions, which
710 overwhelmingly involve exactly the same functions, is to overlook
711 an *explanation* of the construction's formal property and its
712 distribution. Thus, we have seen that the fact that the subject
713 is treated as a unit (so that any auxiliary within the subject
714 is irrelevant) is not mysterious once we recognize that it is a
715 semantic unit. Moreover, the fact that it is the *first* auxiliary of
716 the clause that is inverted is motivated by the functions of the
717 constructions that exhibit SAI.

719 CROSS-LINGUISTIC GENERALIZATIONS 720 ABOUT THE LINKING BETWEEN 721 SEMANTICS AND SYNTAX 722

723 The last type of generalization considered here is perhaps
724 the most straightforward. There are certain claims about
725 how individual semantic arguments are mapped to syntax
726 that have been claimed to require syntactic stipulation, but
727 which follow straightforwardly from the semantic functions of
728 the arguments.

729 Consider the claimed universal that the number of semantic
730 arguments equals the number of overt complements expressed
731 (the “ θ criterion”; see also Lidz et al., 2003a). While the
732 generalization holds, roughly, in English, it does not in many—
733 perhaps the majority—of the world's languages, which readily
734 allow recoverable or irrelevant arguments to be omitted. Even
735 in English, particular constructions circumvent the general
736 tendency. For example, short passives allow the semantic agent
737 or causer argument to be unexpressed (e.g., *The duck was*
738 *killed*), and the “deprofiled object construction” allows certain
739 arguments to be omitted because they are irrelevant (e.g.,
740 *Lions only kill at night*). (Goldberg, 2000). Thus, the original

742 syntactic claim is too strong. A more modest, empirically accurate
743 generalization is captured by the following:

744 Pragmatic Mapping Generalization (Goldberg, 2004):

- 745
746 A) The referents of linguistically expressed arguments are
747 interpreted to be *relevant* to the message being conveyed.
748
749 B) Any semantic participants in the event being conveyed that
750 are *relevant* and *non-recoverable* from context must be overtly
751 indicated.

752 The pragmatic mapping generalization makes use of the fact that
753 language is a means of communication and therefore requires
754 that speakers say as much as is necessary but not more (Paul,
755 1889; Grice, 1975). Note that the pragmatic generation does
756 not make any predictions about semantic arguments that are
757 recoverable or irrelevant. This is important because, as already
758 mentioned, languages and constructions within languages treat
759 those arguments variably.

759 Another general cross-linguistic tendency is suggested by
760 Dowty (1991), who proposed a linking generalization that is
761 now widely cited as capturing the observable (i.e., surface) cross-
762 linguistic universals about how syntactic relations and semantic
763 arguments are linked. Dowty argued that in simple active clauses,
764 *if* there both a subject and an object, *and if* there is an agent-like
765 semantic argument and an undergoer-like semantic argument,
766 then the agent will be expressed by the subject, and the undergoer
767 will be expressed by the direct object (see also Van Valin,
768 1990). Agent-like entities are entities that are volitional, sentient,
769 causal or moving, while undergoers are those arguments that
770 undergo a change of state, are causally affected or are stationary.
771 Dowty further observed that his generalization is violated in
772 syntactically ergative languages, which are quite complicated
773 and do not neatly map the agent-like argument to a subject.
774 In fact, there are no syntactic tests for subjecthood that are
775 consistent across languages so there is no reason to assume
776 that the grammatical relation of subject is universal (Dryer,
777 1997).

778 At the same time, there does exist a more modest “linking”
779 generalization that is accurate: actors and undergoers are
780 generally expressed in prominent syntactic slots (Goldberg,
781 2006). This simpler generalization, which I have called the *salient-*
782 *participants-in-prominent-slots* generalization has the advantage
783 that it accurately predicts that an actor argument without an
784 undergoer, and an undergoer without an actor are also expressed
785 in prominent syntactic positions.

786 The tendency to express salient participants in prominent slots
787 follows from well-documented aspects of our general attentional
788 biases. Humans' attention is naturally drawn to agents, even in
789 non-linguistic tasks. For example, visual attention tends to be
790 centered on the agent in an event (Robertson and Suci, 1980).
791 Speakers also tend to adopt the perspective of the agent of the
792 event (MacWhinney, 1977; Hall et al., 2013). Infants as young as
793 9 months have been shown to attribute intentional behavior even
794 to inanimate objects that have appropriate characteristics (e.g.,
795 motion, apparent goal-directedness) (Csibra et al., 1999). That is,
796 even, pre-linguistic infants attend closely to the characteristics of
797 agents (volition, sentience, and movement) in visual as well as
798 linguistic tasks.

The undergoer in an event is also attention-worthy, as it is generally the endpoint of a real or metaphorical force (Langacker, 1987; Talmy, 1988; Croft, 1991). The tendency to attend closely endpoints of actions that involve a change of state exists even in 6 month old infants (Woodward, 1998), and we know that the effects of actions play a key role in action-representations both in motor control of action and in perception (Prinz, 1990, 1997). For evidence that undergoers are salient in non-linguistic tasks, see also Csibra et al. (1999); Bekkering et al. (2000); Javanovic et al. (2007). For evidence that endpoints or undergoers are salient in linguistic tasks, see Regier and Zheng (2003); Lakusta and Landau (2005), and Lakusta et al. (2007). Thus, the observation that agents and undergoers tend to be expressed in prominent syntactic positions is explained by general facts about human perception and attention.

Other generalizations across languages are also amenable to functional explanations. There is a strong universal tendency for languages to have some sort of construction that can reasonably be termed a “passive.” But these passive constructions only share a general function: they are constructions in which the topic and/or agent argument is essentially “demoted,” appearing optionally or not at all. In this way, passive constructions offer speakers more flexibility in how information is packaged. But whether or which auxiliary appears, whether a given language has one, two, or three passives, whether or not intransitive verbs occur in the pattern, and whether or how the demoted subject argument is marked, all *differ* across different languages (Croft, 2001), and certain languages such as Choctaw do not seem to contain any type of passive (Van Valin, 1980). That is the only robust generalization about passive depends on its function and is very modest: most, but not all languages, have a way to express what is normally the most prominent argument in a less prominent position.

CONCLUSION

When it was first proposed that our knowledge of language was so complex and subtle and that the input was so impoverished that certain syntactic knowledge must be given to us *a priori*, the argument was fairly compelling (Chomsky, 1965). At that time, we did not have access to large corpora of child-directed speech so we did not realize how massively repetitive the input was; nor did we have large corpora of children’s early speech, so we did not appreciate how closely children’s initial productions reflect their input (see e.g., Mintz et al., 2002; Cameron-Faulkner et al., 2003). We also had not yet fully appreciated how statistical learning worked, nor how powerful it was (e.g., Saffran et al., 1996; Gomez and Gerken, 2000; Fiser and Aslin, 2002; Saffran, 2003; Abbot-Smith et al., 2008; Wonnacott et al., 2008; Kam and Newport, 2009). Connectionist and Bayesian modeling had not yet revealed that associative learning and rational inductive inferences could be used to address many aspects of language learning (see e.g., Elman et al., 1996; Perfors et al., 2007; Alishahi and Stevenson, 2008; Bod, 2009). The important role of language’s function as a means of communication was widely ignored (but see e.g., Lakoff, 1969; Bolinger, 1977; DuBois, 1987; Langacker, 1987; Givón, 1991). Finally, the widespread recognition of emergent

phenomena was decades away (e.g., Karmiloff-Smith, 1992; Lander and Schork, 1994; Elman et al., 1996). Today, however, armed with these tools, we are able to avoid the assumption that all languages must be “underlyingly” the same in key respects or learned via some sort of tailor-made “Language Acquisition Device” (Chomsky, 1965). In fact, if Universal Grammar consists only of recursion via “merge,” as Chomsky has proposed (Hauser et al., 2002), it is unclear how it could even begin to address the purported poverty of the input issue in any case (Ambridge et al., 2015).

Humans are unique among animals in the impressive diversity of our communicative systems (Dryer, 1997; Croft, 2001; Tomasello, 2003:1; Haspelmath, 2008; Evans and Levinson, 2009; Everett, 2009). If we assume that all languages share certain important formal parallels “underlyingly” due to a tightly constrained Universal Grammar, except perhaps for some simple parameter settings, it would seem to be an unexplained and maladaptive feature of languages that they involve such rampant superficial variation. In fact, there are cogent arguments against positing innate, syntax-specific, universal knowledge of language, as it is biologically and evolutionarily highly implausible (Christiansen and Kirby, 2003; Chater et al., 2009; Christiansen and Chater, 2016).

Instead, what makes language possible is a certain combination of prerequisites for language, including our pro-social motivation and skill (e.g., Hermann et al., 2007; Tomasello, 2008); the general trade off between economy of effort and maximization of expressive power (e.g., Levy, 2008; Futrell et al., 2015; Kirby et al., 2015; Kurumada and Jaeger, 2015); the power of statistical learning (Saffran et al., 1996; Gomez and Gerken, 2000; Saffran, 2003; Wonnacott et al., 2008; Kam and Newport, 2009); and the fact that frequently used patterns tend to become conventionalized and abbreviated (Heine, 1992; Dabrowska, 2004; Bybee et al., 1997; Verhagen, 2006; Traugott, 2008; Bybee, 2010; Hilpert, 2013; Traugott and Trousdale, 2013; Christiansen and Chater, 2016).

While these prerequisites for language are highly pertinent to the discussion of whether we need to appeal to a Universal Grammar, the present paper has attempted to address a different set of facts. Many generative linguists take the existence of subtle, intricate, knowledge about language that speakers implicitly know without being taught as evidence in favor of the Universal Grammar Hypothesis. By examining certain of these well-studied such cases, we have seen that, while the facts are sometimes even more complex and subtle than is generally appreciated, they do not require that we resort to positing syntactic structures that are unlearned. Instead, these cases are explicable in terms of the *functions of the constructions* involved. That is, the constructionist perspective views intricate and subtle generalizations about language as emerging on the basis of domain-general constraints on perception, attention, and memory, *and* on the basis of the functions of the learned, conventionalized constructions involved. This paper has emphasized the latter point.

Constructionists recognize that languages are not unconstrained in their variation and that various systematic patterns recur in unrelated languages. While certain

913 generalizations follow from domain-general processing
 914 constraints (see e.g., McRae et al., 1998; Hawkins, 1999; Futrell
 915 et al., 2015), this paper as argued that many constraints and
 916 generalizations follow from the functions of the constructions
 917 involved. That is, speakers can combine conventional
 918 constructions in their language on the fly to create new
 919 utterances, but the functions of each of the constructions
 920 involved must be respected. This allows speakers to use language
 921 in dynamic, but delimited ways.

925 REFERENCES

- 926
 927 Abbot-Smith, K., Dittmar, M., and Tomasello, M. (2008). Graded representations
 928 in the acquisition of English and German transitive construction. *Cogn. Dev.*
 929 23, 48–66. doi: 10.1016/j.cogdev.2007.11.002
 930 Alishahi, A., and Stevenson, S. (2008). A computational model of early argument
 931 structure acquisition. *Cogn. Sci.* 32, 789–834. doi: 10.1080/03640210801929287
 932 Ambridge, B., and Goldberg, A. E. (2008). The island status of clausal
 933 complements: evidence in favor of an information structure explanation. *Cogn.*
 934 *Linguist.* 19, 349–381. doi: 10.1515/COGL.2008.014
 935 Ambridge, B., Pine, J. M., and Lieven, E. V. M. (2015). Explanatory adequacy is
 936 not enough: Response to commentators on ‘Child language acquisition: Why
 937 universal grammar doesn’t help’. *Language* 91, e116–e126. doi: 10.1353/lan.
 938 2015.0037
 939 Ambridge, B., Rowland, C. F., Theakston, A. L., and Tomasello, M. (2006).
 940 Comparing different accounts of inversion errors in children’s non-subject wh-
 941 questions: ‘what experimental data can tell us?’ *J. Child Lang.* 33, 519–557. doi:
 942 10.1017/S0305000906007513
 943 Arnold, J. E. (1998). *Reference form and Discourse Patterns*. Doctoral dissertation,
 944 Stanford University.
 945 Auer, P., and Pfänder, S. (eds.). (2011). *Constructions: Emerging and Emergent*.
 946 Berlin: De Gruyter Mouton.
 947 Baker, C. L. (1978). *Introduction to Generative Transformational Syntax*.
 948 Englewood Cliffs, NJ: Prentice-Hall.
 949 Bekkering, H., Wohlschläger, A., and Gattis, M. (2000). Imitation of gestures
 950 in children is goal-directed. *Q. J. Exp. Psychol.* 53A, 153–164. doi:
 951 10.1080/713755872
 952 Blumberg, M. S. (2006). *Basic Instinct: the Genesis of Behavior*. New York, NY:
 953 Avalon Publishing.
 954 Bod, R. (2009). From exemplar to grammar: integrating analogy and probability
 955 learning. *Cogn. Sci.* 13, 311–341.
 956 Bolinger, D. (1977). *Meaning and Form*. London: Longman.
 957 Bresnan, J. W. (1972). *Theory of Complementation in English syntax*. Doctoral
 958 dissertation, Massachusetts Institute of Technology.
 959 Brugman, C., and Lakoff, G. (1987). “The semantics of aux-inversion and anaphora
 960 constraints,” *Proceedings of the Berkeley Linguistic Society* (Berkeley: Berkeley
 961 Linguistics Society).
 962 Bybee, J. (2010). *Language, Usage and Cognition*. Cambridge: Cambridge
 963 University Press.
 964 Bybee, J., Haiman, J., and Thompson, S. (eds.). (1997). *Essays on Language Function
 965 and Language Type*. Amsterdam: Benjamins.
 966 Cameron-Faulkner, T., Lieven, E., and Tomasello, M. (2003). A construction
 967 based analysis of child directed speech. *Cogn. Sci.* 27, 843–873. doi:
 968 10.1207/s15516709cog2706_2
 969 Chafe, W. L. (1987). “Cognitive constraints on information flow,” in *Coherence and
 Grounding in Discourse*, ed R. Tomlin (Amsterdam: Benjamins), 5–25.
 Chafe, W. L. (1994). *Discourse, Consciousness and Time: the Flow and Displacement
 of Conscious Experience in Speaking and Writing*. Chicago, IL: University of
 Chicago Press.
 Chater, N., Reali, F., and Christiansen, M. H. (2009). Restrictions on biological
 adaptation in language evolution. *Proc. Natl. Acad. Sci. U.S.A.* 106, 1015–1020.
 doi: 10.1073/pnas.0807191106
 Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

AUTHOR CONTRIBUTIONS

AG wrote the paper in its entirety with appropriately cited
 references.

ACKNOWLEDGMENTS

I would like to thank Elizabeth Traugott, Jeff Lidz, and Nick
 Enfield for very helpful feedback on an earlier draft of this paper.

- Christiansen, M. H., and Chater, N. (2016). *Creating Language: Integrating
 Evolution, Acquisition, and Processing*. Cambridge, MA: MIT Press.
 Christiansen, M. H., and Kirby, S. (eds.). (2003). *Language evolution*. Oxford:
 Oxford University Press.
 Croft, W. (1991). *Syntactic Categories and Grammatical Relations: The Cognitive
 Organization of Information*. Chicago, IL: University of Chicago Press.
 Croft, W. (2001). *Radical Construction Grammar*. Oxford: Oxford University
 Press.
 Csibra, G., Gergely, G., Biró, S., Koós, O., and Brockbank, M. (1999). Goal-
 attribution and without agency cues: the perception of “pure reason” in infancy.
Cognition 72, 237–267. doi: 10.1016/S0010-0277(99)00039-6
 Culicover, P. W., and Jackendoff, R. (2005). *Simpler Syntax*. Oxford: Oxford
 University Press.
 Dabrowska, E. (2004). *Language, Mind and Brain: Some Psychological and
 Neurological Constraints on Theories of Grammar*. Edinburgh: Edinburgh
 University Press.
 Dale, R. (2003). “One-Anaphora and the case for discourse-driven referring
 expression generation,” in *Proceedings of the Australasian Language Technology
 Workshop* (Melbourne: University of Melbourne).
 Davies, M. (2008). *The Corpus of Contemporary American English (COCA): 400+
 Million Words, 1990-present*. Available online at: <http://www.americancorpus.org>
 Deák, G. (2000). Hunting the fox of word learning: why “Constraints” fail to
 capture it. *Dev. Rev.* 20, 29–80. doi: 10.1006/drev.1999.0494
 Deane, P. (1991). Limits to attention: a cognitive theory of island phenomena.
Cogn. Linguist. 2, 1–63. doi: 10.1515/cogl.1991.2.1.1
 Dowty, D. (1991). Thematic proto-roles and argument selection. *Language* 67,
 547–619. doi: 10.1353/lan.1991.0021
 Dryer, M. (1997). “Are grammatical relations universal?” in *Essays on Language
 Function and Language Type*, J. Bybee, J. Haiman, and S. Thompson
 (Amsterdam: Benjamins), 115–143.
 DuBois, J. W. (1987). The discourse basis of ergativity. *Language* 63, 805–855. doi:
 10.2307/415719
 Elman, J., Bates, E., Johnson, M., Karmiloff-Smith, A., Parisi, D., and Plunkett,
 K. (1996). *Rethinking Innateness: A Connectionist Perspective on Development*.
 Cambridge, MA: MIT Press.
 Engdahl, E. (1997). Relative clause extractions in context. *Work. Pap. Scand. Syntax*
 60, 51–79.
 Erteschik-Shir, N. (1998). “The syntax-focus structure interface,” *Syntax and
 Semantics 29: The Limits of Syntax*, eds P. Culicover and L. McNally (New York,
 NY: Academic Press), 211–240.
 Evans, N., and Levinson, S. C. (2009). The myth of language universals: language
 diversity and its importance for cognitive science. *Behav. Brain Sci.* 32, 429–448.
 doi: 10.1017/S0140525X0999094X
 Everett, D. L. (2009). Pirahã culture and grammar: a response to some criticisms.
Language 85, 405–442. doi: 10.1353/lan.0.0104
 Fillmore, C. J. (1999). “Inversion and constructional inheritance,” in
Lexical and Constructional Aspects of Linguistic Explanation, eds G.
 Webelhuth, P. J. Koenig, and A. Kathol (Stanford, CA: CSLI Publication),
 118–129.
 Fiser, J., and Aslin, R. N. (2002). Statistical learning of new visual feature
 combinations by infants. *Proc. Natl. Acad. Sci. U.S.A.* 99, 15822–15826. doi:
 10.1073/pnas.232472899

- 1027 Fodor, J., Bever, A., and Garrett, T. G. (1974). *The Psychology of Language: An*
1028 *Introduction to Psycholinguistics and Generative Grammar*. New York, NY:
1029 McGraw-Hill.
- 1030 Futrell, R., Mahowald, K., and Gibson, E. (2015). Large-scale evidence of
1031 dependency length minimization in 37 languages (vol 112, pg 10336, 2015).
1032 *Proc. Natl. Acad. Sci. U.S.A.* 112, E5443–E5444. doi: 10.1073/pnas.1502134112
- 1033 Givón, T. (1991). “Serial verbs and the mental reality of ‘event’: grammatical vs.
1034 cognitive packaging,” in *Approaches to Grammaticalization*, Vol. 1, eds E. C.
1035 Traugott and B. Heine (Amsterdam: John Benjamins Publishing Company),
1036 81–127.
- 1037 Givón, T. (2001). *Syntax: An Introduction, Vol. 1*. Amsterdam: John Benjamins
1038 Publishing.
- 1039 Goldberg, A. E. (2000). Patient arguments of causative verbs can be omitted: the
1040 role of information structure in argument distribution. *Lang. Sci.* 34, 503–524.
1041 doi: 10.1016/S0388-0001(00)00034-6
- 1042 Goldberg, A. E. (2004). But do we need universal grammar? Comment on Lidz
1043 et al. (2003). *Cognition* 94, 77–84. doi: 10.1016/j.cognition.2004.03.003
- 1044 Goldberg, A. E. (2006). *Constructions at Work: The Nature of Generalization in*
1045 *Language*. Oxford: Oxford University Press.
- 1046 Goldberg, A. E. (2009). The nature of generalization in language. *Cogn. Linguist.*
1047 20, 93–127. doi: 10.1515/COGL.2009.005
- 1048 Goldberg, A. E. (2013). “Backgrounded constituents cannot be “extracted,”” in
1049 *Experimental Syntax and Island Effects*, ed J. Sproule (Cambridge: Cambridge
1050 University Press), 221–238. doi: 10.1017/CBO9781139035309.012
- 1051 Goldberg, A. E., and de Giudice, A. D. (2005). Subject-auxiliary inversion: a natural
1052 category. *Linguist. Rev.* 22, 411–428. doi: 10.1515/tlir.2005.22.2-4.411
- 1053 Goldberg, A. E., and Michaelis, L. A. (2015). One from many: anaphoric one and
1054 its relationship to numeral one. *Cogn. Sci.*
- 1055 Gomez, R. L., and Gerken, L. D. (2000). Infant artificial language learning
1056 and language acquisition. *Trends Cogn. Sci.* 4, 178–186. doi: 10.1016/S1364-
1057 6613(00)01467-4
- 1058 Gregory, M. L., and Michaelis, L. A. (2001). Topicalization and left dislocation: a
1059 functional opposition revisited. *J. Pragmat.* 33, 1665–1706. doi: 10.1016/S0378-
1060 2166(00)00063-1
- 1061 Grice, H. P. (1975). “Logic and conversation,” in *Syntax and Semantics, Vol. 3:*
1062 *Speech Acts*, eds P. Cole and J. Morgan (New York, NY: Academic Press), 41–58.
- 1063 Hall, M. L., Mayberry, R. I., and Ferreira, V. S. (2013). Cognitive constraints on
1064 constituent order: evidence from elicited pantomime. *Cognition* 129, 1–17. doi:
1065 10.1016/j.cognition.2013.05.004
- 1066 Halliday, M. A. (1967). Notes on transitivity and theme in English Part, I.
1067 *J. Linguist.* 3, 37–81. doi: 10.1017/S0022226700012949
- 1068 Haspelmath, M. (2008). “Parametric versus functional explanation of syntactic
1069 universals,” in *The Limits of Syntactic Variation*, ed T. Biberauer (Amsterdam:
1070 Benjamins), 75–107.
- 1071 Hauser, M. D., Chomsky, N., and Fitch, W. T. (2002). The faculty of language:
1072 what is it, who has it, and how did it evolve? *Science* 298, 1569–1579. doi:
1073 10.1126/science.298.5598.1569
- 1074 Hawkins, J. A. (1999). Processing complexity and filler-gap dependencies across
1075 grammars. *Language* 75, 244–285. doi: 10.2307/417261
- 1076 Heine, B. (1992). Grammaticalization chains. *Stud. Lang.* 16, 335–368. doi:
1077 10.1075/sl.16.2.05hei
- 1078 Hermann, E., Call, J., Hernández-lloreda, M. V., Hare, B., and Tomasello,
1079 M. (2007). Humans have evolved specialized skills of social cognition: The
1080 cultural intelligence hypothesis. *Science* 317, 1360–1366. doi: 10.1126/science.
1081 1146282
- 1082 Hicks, G. (2003). “So Easy to Look At, So Hard to Define”: *Tough Movement in the*
1083 *Minimalist Framework*. Doctoral dissertation, University of York.
- 1084 Hilpert, M. (2013). *Constructional Change in English: Developments in Allomorphy,*
1085 *Word-Formation and Syntax*. Cambridge: Cambridge University Press.
- 1086 Hopper, P. J. (1987). “Emergent grammar,” in *Berkeley Linguistics Society 13:*
1087 *General Session and Parasession on Grammar and Cognition*, eds J. Aske,
1088 N. Berry, L. Michaelis, and H. Filip (Berkeley: Berkeley Linguistics Society),
1089 139–157.
- 1090 Hornstein, N., and Lightfoot, D. (1981). “Introduction,” in *Explanation in*
1091 *Linguistics: The Logical Problem of Language Acquisition*, eds N. Hornstein and
1092 D. Lightfoot (London: Longman), 9–31.
- 1093 Jackendoff, R. (1977). *X-Bar Syntax: A Study of Phrase Structure*. Cambridge, MA:
1094 MIT Press.
- 1095 Javanovic, B., Kiraly, I., Elsner, B., Gergely, G., Prinz, W., Aschersleben, G., et al.
1096 (2007). The role of effects for infants’ perception of action goals. *Psychologia* 50,
1097 273–290. doi: 10.2117/psychoc.2007.273
- 1098 Kam, X. N. C., Stoyaneshka, I., Torniyova, L., Fodor, J. D., and Sakas, W. G.
1099 (2008). Bigrams and the richness of the stimulus. *Cogn. Sci.* 32, 771–787. doi:
1100 10.1080/03640210802067053
- 1101 Kam, C. L., and Newport, E. L. (2009). Getting it right by getting it
1102 wrong: when learners change languages. *Cogn. Psychol.* 59, 30–66. doi:
1103 10.1016/j.cogpsych.2009.01.001
- 1104 Karmiloff-Smith, A. (1992). *Beyond Modularity: A Developmental Perspective on*
1105 *Cognitive Science*. Cambridge, MA: MIT Press.
- 1106 Karmiloff-Smith, A. (2006). The tortuous route from genes to behavior: a
1107 neuroconstructivist approach. *Cogn. Affect. Behav. Neurosci.* 6, 9–17. doi:
1108 10.3758/CABN.6.1.9
- 1109 Kirby, S. (2000). “Syntax without natural selection: how compositionality emerges
1110 from vocabulary in a population of learners,” in *The Evolutionary Emergence*
1111 *of Language: Social Function and the Origins of Linguistic Form*, ed C. Knight
1112 (Cambridge: Cambridge University Press), 1–19.
- 1113 Kirby, S., Tamariz, M., Cornish, H., and Smith, K. (2015). Compression and
1114 communication in the cultural evolution of linguistic structure. *Cognition* 141,
1115 87–102. doi: 10.1016/j.cognition.2015.03.016
- 1116 Kuno, S. (1976). “Subject, theme, and the speaker’s empathy: a reexamination of
1117 relativization phenomena in subject and topic,” in *Subject and Topic*, ed Li (New
1118 York, NY: Academic Press), 417–444.
- 1119 Kurumada, C., and Jaeger, T. F. (2015). Communicative efficiency in language
1120 production: optional case-marking in Japanese. *J. Mem. Lang.* 83, 152–178. doi:
1121 10.1016/j.jml.2015.03.003
- 1122 Labov, W. (1968). *A Study of the Non-Standard English of Negro and Puerto Rican*
1123 *Speakers in New York City. Volume I: Phonological and Grammatical Analysis.*
1124 Final Report of Education Cooperative Research Project.
- 1125 Lakoff, G. (1969). *On Generative Semantics*. Bloomington, IN: Indiana University
1126 Linguistics Club.
- 1127 Lakoff, G. (1970). “Pronominalization, negation and the analysis of adverbs,”
1128 in *Readings in English Transformational Grammar*, eds R. Jacobs, and P.
1129 Rosenbaum (Waltham, MA: Ginn and Co), 145–165.
- 1130 Lakusta, L., and Landau, B. (2005). Starting at the end: the importance of
1131 goals in spatial language. *Cognition* 96, 1–33. doi: 10.1016/j.cognition.2004.
1132 03.009
- 1133 Lakusta, L., Wagner, L., O’Hearn, K., and Landau, B. (2007). Conceptual
1134 foundations of spatial language: evidence for a goal bias in infants. *Lang. Learn.*
1135 3, 179–197. doi: 10.1080/15475440701360168
- 1136 Lambrecht, K. (1994). *Information Structure and Sentence Form: A Theory of*
1137 *Topic, Focus, and the Mental Representations of Discourse Referents*. Cambridge:
1138 Cambridge University Press.
- 1139 Lander, E. S., and Schork, N. J. (1994). Genetic dissection of complex traits. *Science*
1140 265, 2037–2048.
- 1141 Langacker, R. W. (1987). *Foundations of Cognitive Grammar Volume I*. Stanford,
1142 CA: Stanford University Press.
- 1143 Langacker, R. W. (1991). *Foundations of Cognitive Grammar Volume 2: Descriptive*
1144 *Applications*. Stanford, CA: Stanford University Press.
- 1145 Langacker, R. W. (2012). *How to Build an English Clause*. San Diego, CA:
1146 University of California.
- 1147 Levy, R. (2008). “A noisy-channel model of rational human sentence
1148 comprehension under uncertain input,” in *Conference on Empirical Methods in*
1149 *Natural Language Processing* (Honolulu, HI).
- 1150 Lewis, J., and Elman, J. (2001). “A connectionist investigation of linguistic
1151 universals: learning the unlearnable,” in *Proceedings of the Fifth International*
1152 *Conference on Cognitive and Neural Systems, Center for Adaptive Systems and*
1153 *the Department of Cognitive and Neural Systems*, Boston University, eds R.
1154 Amos, C. Bradford, C. Jefferson, and D. Meyers (Boston, MA).
- 1155 Lidz, J., Gleitman, H., and Gleitman, L. (2003a). Understanding how input matters:
1156 verb learning and the footprint of universal grammar. *Cognition* 87, 151–178.
1157 doi: 10.1016/S0010-0277(02)00230-5
- 1158 Lidz, J., Waxman, S., and Freedman, J. (2003b). What infants know about syntax
1159 but couldn’t have learned: experimental evidence for syntactic structure at 18
1160 months. *Cognition* 89, 295–303. doi: 10.1016/S0010-0277(03)00116-1
- 1161 MacWhinney, B. (1977). Starting points. *Language (Baltim.)* 53, 152–168. doi:
1162 10.2307/413059

- 1141 McRae, K., Spivey-Knowlton, M. J., and Tanenhaus, M. K. (1998). Modeling
1142 the influence of thematic fit (and other constraints) in on-line sentence
1143 comprehension. *J. Mem. Lang.* 38, 283–312. doi: 10.1006/jmla.1997.2543
- 1144 Merchant, J. (2009). Phrasal and clausal comparatives in Greek and the
1145 abstractness of syntax. *J. Greek Linguist.* 9, 134–164. doi: 10.1163/156658409
X12500896406005
- 1146 Michaelis, L. A., and Lambrecht, K. (1996). Toward a construction-based model
1147 of language function: the case of nominal extraposition. *Language (Baltim.)* 72,
1148 215–247. doi: 10.2307/416650
- 1149 Michaelis, L., and Francis, H. (2007). “Lexical subjects and the conflation strategy,”
1150 *The Grammar Pragmatics Interface: Essays in Honor of Jeanette K. Gundel, Vol.*
1151 *155*, eds N. Hedberg and R. Zacharski (Stanford, CA: Center for the Study of
1152 Language and Information Publishing), 19.
- 1153 Mintz, T. H., Newport, E. L., and Bever, T. G. (2002). The distributional structure
1154 of grammatical categories in speech to young children. *Cogn. Sci.* 26, 393–424.
1155 doi: 10.1207/s15516709cog2604_1
- 1156 Morgan, J. L. (1975). Some interactions of syntax and pragmatics. *Syntax and*
1157 *Semantics: Speech Acts*, vol. 3, ed. by P. Gole and J.L. Morgan. New York:
1158 Academic Press.
- 1159 Nevalainen, T. (1997). Recycling inversion: the case of initial adverbs and negators
1160 in early modern english. *Studia Anglica Posnaniensia* 31, 202–214.
- 1161 Paul, H. (1889). *Principles of the History of Language*. Transl. by H. A. Strong. New
1162 York, NY: MacMillan and Company.
- 1163 Payne, J., Pullum, G. K., Scholz, B. C., and Berlage, E. (2013). Anaphoric *one* and
1164 its implications. *Language* 89, 794–829. doi: 10.1353/lan.2013.0071
- 1165 Pearl, L., and Sprouse, J. (2013). Syntactic islands and learning biases:
1166 combining experimental syntax and computational modeling to
1167 investigate the language acquisition problem. *Lang. Acquis.* 20, 23–68.
1168 doi: 10.1080/10489223.2012.738742
- 1169 Perfors, A., Kemp, C., Tenenbaum, J., and Wonnacott, E. (2007). *Learning*
1170 *Inductive Constraints: The Acquisition of Verb Argument Constructions*.
1171 University College London.
- 1172 Perfors, A., Tenenbaum, J. B., and Regier, T. (2011). The learnability
1173 of abstract syntactic principles. *Cognition* 118, 306–338. doi:
1174 10.1016/j.cognition.2010.11.001
- 1175 Polinsky, M. (1998). “A non-syntactic account of some asymmetries in the double
1176 object construction,” in *Conceptual Structure and Language: Bridging the Gap*,
1177 ed J.-P. Koenig (Stanford, CA: CSLI), 98–112.
- 1178 Postal, P., and Ross, J. R. (1971). Tough movement Si, tough deletion No! *Linguistic*
1179 *2*, 544–546.
- 1180 Prinz, W. (1990). “A common coding approach to perception and action,” in
1181 *Relationships Between Perception and Action: Current Approaches*, O. Neumann
1182 and W. Prinz (Berlin: Springer Verlag), 167–201.
- 1183 Prinz, W. (1997). Perception and action planning. *Eur. J. Cogn. Psychol.* 9, 129–154.
1184 doi: 10.1080/713752551
- 1185 Pullum, G. K., and Scholz, B. C. (2002). Empirical assessment of stimulus poverty
1186 arguments. *Linguistic Rev.* 19, 9–50. doi: 10.1515/tlir.19.1-2.9
- 1187 Radford, A. (1988). *Transformational Grammar: A First Course*. Cambridge:
1188 Cambridge University Press. doi: 10.1017/CBO9780511840425
- 1189 Reali, F., and Christiansen, M. H. (2005). Uncovering the richness of the
1190 stimulus: structure dependence and indirect statistical evidence. *Cogn. Sci.* 29,
1191 1007–1028. doi: 10.1207/s15516709cog0000_28
- 1192 Regier, T., and Zheng, M. (2003). “An attentional constraint on spatial meaning,”
1193 in *Proceedings of the 25th Annual Meeting of the Cognitive Science Society*
1194 (Boston, MA).
- 1195 Robertson, S. S., and Suci, G. J. (1980). Event perception in children in early stages
1196 of language production. *Child Dev.* 51, 89–96. doi: 10.2307/1129594
- 1197 Ross, J. R. (1967). *Infinite Syntax*. New York, NY: Ablex Publishing.
- 1198 Rowland, C. F. (2007). Explaining errors in children’s questions. *Cognition* 104,
1199 106–134. doi: 10.1016/j.cognition.2006.05.011
- 1200 Saffran, J. R. (2003). Statistical language learning: mechanisms and constraints.
1201 *Curr. Dir. Psychol. Sci.* 12, 110–114. doi: 10.1111/1467-8721.01243
- 1202 Saffran, J. R., Aslin, R. N., and Newport, E. L. (1996). Statistical learning by
1203 8-month-old infants. *Science* 274, 1926–1928.
- 1204 Sag, I. A. (2010). English filler-gap constructions. *Language* 86, 486–545. doi:
1205 10.1353/lan.2010.0002
- 1206 Sprouse, J., and Hornstein, N. (eds.). (2013). *Experimental Syntax and*
1207 *Island Effects*. Cambridge, UK: Cambridge University Press. doi:
1208 10.1017/cbo9781139035309
- 1209 Strawson, P. F. (1964). Identifying reference and truth—values. *Theoria* 30, 96–118.
1210 doi: 10.1111/j.1755-2567.1964.tb00404.x
- 1211 Takami, K.-I. (1989). Preposition stranding: arguments against syntactic analyses
1212 and an alternative functional explanation. *Lingua* 76, 299–335.
- 1213 Talmy, L. (1988). Force dynamics in language and cognition. *Cogn. Sci.* 12, 49–100.
1214 doi: 10.1207/s15516709cog1201_2
- 1215 Tomasello, M. (2003). *Constructing a Language: A Usage-Based Theory of Language*
1216 *Acquisition*. Boston, MA: Harvard University Press.
- 1217 Tomasello, M. (2008). *Origins of Human Communication*. Cambridge: MIT Press.
- 1218 Traugott, E. C. (2008). “All that he endeavoured to prove was..”: on the emergence
1219 of grammatical constructions in dialogal and dialogic contexts,” in *Language*
1220 *in Flux: Dialogue Coordination, Language Variation, Change and Evolution*, eds
1221 R. Cooper and R. Kempson (London: Kings College Publications), 143–177.
- 1222 Traugott, E. C., and Trousdale, G. (2013). *Constructionalization and*
1223 *Constructional Changes*. Oxford: Oxford University Press. doi:
1224 10.1093/acprof:oso/9780199679898.001.0001
- 1225 Van Valin, R. D. Jr. (1980). On the distribution of passive and antipassive
1226 constructions in universal grammar. *Lingua* 50, 303–327. doi: 10.1016/0024-
1227 3841(80)90088-1
- 1228 Van Valin, R. D. Jr. (1990). Semantic parameters of split intransitivity. *Language*
1229 66, 212–260. doi: 10.2307/414886
- 1230 Van Valin, R. D. Jr. (1998). “The acquisition of Wh-questions and the mechanisms
1231 of language acquisition,” in *The New Psychology of Language: Cognitive and*
1232 *Functional Approaches to Language Structure*, ed M. Tomasello (Hillsdale, NJ:
1233 Lawrence Erlbaum), 137–152.
- 1234 Verhagen, A. (2006). “On subjectivity and ‘long distance Wh-movement,’” in
1235 *Subjectification: Various Paths to Subjectivity*, eds A. Athanasiadou, C. Canakis
1236 and B. Cornilli (Berlin; New York, NY: Mouton de Gruyter), 323–346.
- 1237 Wonnacott, E., Newport, E. L., and Tanenhaus, M. K. (2008). Acquiring
1238 and processing verb argument structure: distributional learning in a
1239 miniature language. *Cogn. Psychol.* 56, 165–209. doi: 10.1016/j.cogpsych.2007.
1240 04.002
- 1241 Woodward, A. L. (1998). Infants selectively encode the goal object of an actor’s
1242 reach. *Cognition* 69, 1–34. doi: 10.1016/S0010-0277(98)00058-4

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

Copyright © 2016 Goldberg. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.