CAMBRIDGE
UNIVERSITY PRESS

ARTICLE

Constructivist Approaches to First Language Acquisition

Heike BEHRENS

University of Basel
Address for correspondence: Heike Behrens, University of Basel - German Department, Nadelberg 4 Basel
4001, Switzerland. E-mail: Heike.Behrens@unibas.ch

Abstract

Constructivist approaches to language acquisition predict that form-function mappings are derived from distributional patterns in the input, and their contextual embedding. This requires a detailed analysis of the input, and the integration of information from different contingencies. Regarding the acquisition of morphology, it is shown which types of information leads to the induction of (lexical) categories, and to paradigm building. Regarding the acquisition of word order, it is shown how languages with fixed or variable word order profit from stable syntactic hyperschemas, but require a more detailed analyses of the form-function contingencies to identify the underlying, more specific semantic, syntactic and morphological patterns. At a theoretical level, it is shown how findings from acquisition and processing converge into new linguistic theories that aim to account for regular as well as irregular phenomena in language.

Definition and History of Constructivism

Constructivism sees learning as a process by which the learners construct new knowledge based on their current knowledge by integrating the incoming information. With respect to language acquisition, it is conceptually related to acquisition theories such as *emergentism* (with a focus on how new qualities/ representations arise from the interaction of incoming and existing information, for example, how linguistic categories are constructed from distributional information, and *usage-based theories* which focus on the homology between the information attested in the input and the child's developing language use. The focus is on the child's ability to generalize over the input by studying the affordances of the input through targeted corpus analyses of Child Directed Speech; the way in which the child extracts information and generalizes it to new, related situations; and the generalization processes as such, like the speed and nature of schema and category formation, in dependence of the frequency and functional transparency of the input.

CrossMark

Jerome Bruner, the person associated most prominently with formulating constructivist approach to language learning, sketched the task of the language learner as threefold (Bruner, 1983): to learn the linguistic system (syntax), to learn to refer and express meaning (semantics), and to get something done with words and communicate effectively (pragmatics). While the acquisition of morphosyntax is at the heart of this volume, constructivist and usage-based theories start out from this comprehensive view of language use, and see language learning as driven by the need to communicate. This has implications for what has to be explained (language as *explanandum*, see below), as well as for the resources available to the learner (information about the linguistic system itself and the distribution and function of units in the input, the *explanans*). Current versions of constructivism are closely aligned with connectionism (probability-based pattern association) and Construction Grammar as a linguistic theory.

Constructivist theories explain how children reconstruct the system of their target language(s) given their emergent social-cognitive abilities, and their growing experience with the ambient language usage. They start out from the assumption that there innate no innate specifically linguistic knowledge in the form of a "Universal Grammar" that serves as a blueprint of all possible grammatical sentences of all languages, but that children derive their knowledge of language from their environment or input. Language and cognition are dynamic: Language use affects learning because different input styles will lead to individual differences in the learning outcome, just as subsequent changes in language use will affect the linguistic system (language change, lectal variation) (Bybee, 2010; Diessel, 2019).

A major branch of constructivist theories subscribes to the usage-based approach in linguistics. The term was coined in the 1980ies by Ronald Langacker (1987) and applied to acquisition most notably by Michael Tomasello:

> In usage based models of language … all things flow from the actual usage events in which people communicate linguistically with one another. The linguistic skills that a person possesses at any given moment in time - in the form of a "structured inventory of symbolic units" - result from her accumulated experience with language across the totality of usage events in her life. (Tomasello, 2000, p. 61–62)

This leads to the following general predictions:

a) Usage-based theory is compatible with universal linguistic categories, but these would be derived from general processing constraints because only structures that are learnable by the next generation can be transmitted (Christiansen & Chater, 2016; Greenhill et al., 2017; Ibbotson, 2020; Kirby, 2017). Hence, linguistic structures are shaped by domain general learning processes (Ibbotson, 2020).

b) Constructivist research covers the whole range of linguistic diversity. Individual languages pose individual acquisition spaces (Slobin, 1985), such that children induce the form-function patterns of their target language from the input. Hence, we not only expect typological differences between vastly different languages, but also intra-typological differences between closely related languages or varieties of a language (e.g., British vs. America English). Consequently, there is a close match between the linguistic structures used by the child and the input s/he receives, which is particularly noticeable in intra-typological comparisons.

c) If linguistic knowledge is language-specific, emergent, and usage-based, different experiences should lead to individual differences in linguistic representations (Dąbrowska, 2012; Kidd, Donnelly & Christiansen, 2018).

Linguistic Theory: Cognitive Linguistics and Construction Grammar

Construction Grammar integrates formal and functional properties of regular as well as (partially) idiomatic constructions (Goldberg, 1995; Tomasello, 1998). Tomasello furthermore draws on elements of Langacker's Cognitive Grammar (1987) regarding the nature of linguistic representation and generalization. The view on language and cognition is broader than that of linguistic nativism with its focus on (innate) syntax, as the task of the learner is to acquire *all* of language, irrespective of how common or generalizable certain constructions are (Allen & Behrens, 2019; Ambridge & Lieven, 2011; Ibbotson, 2013, 2020).

The Target of language development: The nature of linguistic representations

In order to define acquisition processes, not only the *explanans* (learning mechanisms, available data) needs to be identified, but also the *explanandum* (theory of language). Usage-based implementations of Construction Grammar challenge the assumption that language is in essence a structured rule-based system, and propose that it is an inventory of more or less variable schemas instead:

> (W)hen speakers use language for communicative purposes, they repeat recurrent linguistic forms to reach recurrent communicative goals. The repeated correlations between linguistic forms and communicative goals contribute to keeping two kinds of processes in motion: social processes taking place in speech communities which establish and sustain linguistic conventions, and cognitive processes operating in the minds of individual speakers which establish and sustain individual linguistic knowledge. usage and the social and cognitive processes keep reinforcing each other as long as speakers keep using the language, and this allows shared conventions and individual knowledge to emerge, persist, change, and embrace more or less creative innovations, both in language and in the external world. All this taken together, and in fact not much else, makes a language. (Schmid, 2020, p. 3)

Language emerges through entrenchment in the individual mind, and conventionalization as a cooperative process in the society. Current usage-based models see linguistic structures as an emergent property of communication and routinization. Shared communicative practices lead to the stabilization of structures, which are flexible at the same time because communicative conventions are subject to change. As differences in the generalizability of constructions are the norm, low-frequency or emergent schemas are a part of language just as the more regular ones. Linguistic "rules" are very general schemas (Langacker, 1987), but do not hold a special status in the system (cf. the "rule-list-fallacy", Langacker, 1987, p. 42).

Individual differences

Constructivist theories as discussed here do not assume innate linguistic representations shared by all speakers. This alters the task of language learning from activating innate

settings and linking them to the units of the E-language, to reconstructing the underlying system from the evidence learners hear:

> From a usage-based perspective, such (individual) differences are expected. Language learners do not have direct access to the cognitive routines that other speakers use to produce utterances, only to the utterances themselves. (Dąbrowska, 2020, p. 215)

Speakers converge by and large to the certain linguistic conventions, but this does not entail that they share the same mental representation (Dąbrowska, 2020; Kidd & Donnelly, 2020; Kidd et al., 2018). Utterance can be (re)produced by different cognitive routines, given different types of experiences of the speakers, and different kinds or degrees of generalization. Dąbrowska (2020) calls this *multicollinearity*: The convergence of cues can lead to convergence between speakers, even if they identify and process different cues.

A further implication of individual differences is that speakers' generalizations may differ from the generalizations or "rules" that linguists identify (Dąbrowska, 2020). Valid generalizations about aspects of a language need not be "captured in the mental grammars of all, or even most, speakers" of that language (p. 223). L2-learners of Polish genitives but also native speakers come up with different generalization patterns in task were they have to inflect unknown words. While some pick up on the semantics underlying the distribution of allomorphs (the ending -*a* correlates highly with animate masculine nouns, whereas -*u* correlates with inanimate nouns), others seem to have rote-learned the markers and inflect new nouns based on phonological similarities, or even randomly.

The continuity, linking and developmental problem of acquisition

Recall that the rationale for proposing a Universal Grammar was that certain linguistic operations could not be abstracted from language use, because they relied on structure-dependent operations which are conceived as symbolic operations. A further assumption is that innate categories allow the child to operate with adult-like categories (continuity assumption). According to Pinker (1984) continuity in representation helps to avoid the problem of defining transitory categories and how/when they turn into adult-like categories. If one assumes that the mind operates with symbolic rules, a form of "mentalese" (Pinker, 1994), continuity is needed. In this model, the mind translates the items to be processed (e.g., sequences of morphemes) into grammatical categories that the syntactic rules operate over. Such a "dual-mechanism model" reduces storage or memory load, since only exceptions to the rules (e.g., irregular forms) need to be memorized like lexical items, whereas all regular structures would be generated by rule (Pinker & Prince, 1991). This model requires that all incoming linguistic strings can be parsed and linked unambiguously to their respective grammatical categories in order for the rules to work (the "linking-problem"), as well as a specification of all the operations units of a certain category can undergo in the operation of "feature checking" in the merging process as proposed in the recent minimalist program (Yang, Crain, Berwick, Chomsky & Bolhuis, 2017).

These mechanisms are not necessary if one does not assume innate linguistic categories, but if more or less generalizable schemas are derived from input patterns.

The alternative "single mechanism model" was first implemented in connectionist models and based on findings from cognitive neuroscience (Elman, Bates, Johnson, Karmiloff-Smith, Parisi & Plunkett, 1996; Rumelhart & McClelland, 1987; Twomey & Westermann, 2019). The strength of constructivist accounts is their focus on productivity in order to determine the nature and extent of the generalizations children make across development by looking at detailed naturalistic corpora, or by experimentally testing children's performance with familiar and less familiar items. This makes the single-mechanism model parsimonious, as models that can learn structures with low generalizability should easily learn regular phenomena as well (Fillmore, 1988).

Language learning as a re-construction of the target language

Like connectionist models, usage-based approaches to acquisition are based on pattern association. For the beginning language learner, some of the linguistically complex units are represented as chunks with no representation of the internal constituency. Automatized "units" facilitate processing:

> Unit: A cognitive structure mastered by a speaker to the point that it can be employed in a largely automated fashion, without requiring attention to its individual parts or their arrangements. A unit is sufficiently well entrenched to be easily evoked as an integrated whole, that is, carried out more or less automatically once initiated. (Langacker, 1987, p. 494)

Consequently, constructivist accounts ask: (a) what are the units developing learners operate with, and why; and (b) how do they generalize over different units? A key finding is that children's early generalizations are item-based or low-level schemas. E.g., children first learn about the participants in events of kissing (*kisser – kissee*), throwing (*thrower – thing-thrown – recipient*), before they see semantically based similarities (agent, patient, recipient, instrument etc.). Later, such roles can become associated with domain-general grammatical functions like subject and object (Tomasello, 1992). A fundamental process is drawing associations across related instances. In order to do so, children need to build memory traces that allow them to compare new instances with entrenched ones:

> This accumulated linguistic experience undergoes **processes of entrenchment**, due to **repeated uses** of particular expressions across usage events, and **abstraction**, due to **type variation** in constituents of particular expressions across usage events. (Tomasello, 2000, p. 62).

Langacker (2000, p. 4–6) defines the processes of schema formation and categorization for linguistic units. For example, learners recognize that reference to plural entities or past events correlate with certain affixes. These can develop into schematic frames like *she VERB-ed* or *two NOUN-s*. If learners already knows a few past tense forms, and hears a new form like *gorped* in a similar context, they will be able to associate or categorize this new form with known past tense forms. In sum, constructivist accounts rely on associative learning mechanism (as in connectionist models). But the human learner – unlike current computational models – has full access to the information available through the communicative grounding, and s/he has a

communicative intention or goal. Hence, associative learning mechanisms like pattern finding *and* social-cognitive learning mechanisms like intention reading work together in establishing form-function relationships (Tomasello, 2003).

Complexity reduction through input simplification

The learning mechanisms proposed here draw on the child's ability to distil categorical and functional information from input patterns. For a realistic assessment of the input conditions, detailed analyses of the distributional evidence available to the child in (dense) corpora of naturalistic speech are made (Theakston & Lieven, 2017). There is robust evidence that Child Directed Speech (CDS) represents only a subset of the complexity and variation attested in adult language. In a comparative study on Dutch, Danish, German and Hebrew, Ravid et al. (2008) found that CDS typically uses a smaller core set of markers which depict certain phonotactic regularities more clearly than the more variable adult–adult language. Xanthos et al. (2011) demonstrated complexity reduction in terms of smaller paradigms even in morphologically rich languages. Allen (2017) makes a similar point for input in polysynthetic Inuit languages. Caregivers use a simplified register including shorter nursery words to younger children. Complexity reduction in the input of course increases the learnability. Hence, distributional analyses of CDS are essential to derive corpus-based control-data for experimental studies.

Predictions for the acquisition of specific domains

In the following I will characterize the predictions for morphological categories, and then deal more specifically with the rich set of factors that contribute to the acquisition of fixed and variable word order constructions.

A. Acquisition of morphology

Learning morphology involves two problems:

   a) *Paradigm-building*: Identifying the inflectional categories encoded in a language, and the markers in different distributional contexts.
   b) *Rule-abstraction*: identifying the criteria for allomorphy, i.e., when one inflectional category is encoded by several markers (e.g., regular vs. irregular markers).

Rule-abstraction has been subject to fierce and detailed theoretical debates between association learning (Kirov & Cotterell, 2018) and rule-based dual-mechanisms accounts (Pinker & Prince, 1991), and will not be discussed here. Less attention has been paid to paradigm-building and the nature of morphological categories. Here, the child has to identify the compositional nature of word forms as well as the functions that the affixes encode. In a large typological survey, Slobin (1997) shows that there is no fixed set of privileged grammaticizable notions, so learning mechanisms are needed that can account for the full range of language-specific categories and their idiosyncrasies.

   The paradigm structure of morphological categories can be quite different. Many languages encode subject-verb agreement, but only a few have a specific marker for each person-number combination (e.g., Latin), whereas others have some syncretism

(the same form for several person-number distinctions) like German where 1st and 3rd person plural endings are identical. Yet other languages have broken paradigms with overt marking in only a subset of the paradigm, as in English, where only third person singular present tense is overtly encoded, but not on modal verbs. Such isolated inflectional markers are often remnants of earlier, more complete paradigms where the high frequency of some forms has led to conserving effects and occasional redundant marking (Bybee, 2010), as subject-verb relations is now reliably encoded by word order in English. As a consequence, a learning mind that looks for specific, innate categories and/or strives for maximal generalization would be confronted with a scenario where not all categories that might be expected (e.g., number or tense) are in fact realized in a given language, while the same language might encode more particular, non-universal categories that necessarily have to be derived from language use, and where the paradigms are often not only irregular but broken. Therefore morphology provides a parade-case where irregularity and language-specificity is the norm rather than the exception. So what are the cues that children can use to learn morphology? The following predictions should apply to all morphological categories, irrespective of how regular or productive they are.

Segmentation of morphemes through variation

The first task is to identify the compositional nature of word forms. In English, this is an additive process as children tend to start out with uninflected forms and acquire the rather limited number of inflectional and derivational morphemes in a rather slow fashion (Brown, 1973). More commonly are deconstructive processes in languages where the child only encounters inflected word forms, but no bare stems. Here, tracing the formal variation in dependence of the context is needed. This process is both word-based and pattern-based. Regarding individual words (e.g., nouns or verbs), children are informed by changes of the word form in different distributional contexts that allows them to segment words into stem and affix(es) and build (mini-) paradigms (Ravid, Keuleers & Dressler, 2020).

Semantic accessibility

A further factor is semantic accessibility. Plural inflection is typically acquired early because it has a clear observable correlate (Clark & Nikitina, 2009), but this is not the case for all morphological categories. The classification of German nouns into masculine, feminine and neuter is not based on biological sex, but mainly based on phonotactic properties of the noun, with some semantic predictors (Szagun, Stumper, Sondag & Franik, 2007). The different inflectional classes of the Polish genitive have some semantic correlates, but they are not identified by all speakers (Dąbrowska, 2012).

Prediction regarding paradigms with different phonological or semantic family size

In cases where the inflectional distinctions do not correlate with semantic categorization, forms will have to be stored as exemplars with low generalizability, unless phonological cues will allow pattern abstraction. The prediction that therefore follows here is an interaction between form frequency and phonological neighborhood density on the proportion of correctly-inflected forms vs errors (e.g., in an elicited production study of verb tense and agreement morphology), such that higher form frequency predicts

higher rates of correct production (and fewer errors) unless the phonological neighborhood density is high. Repeated failure to find this interaction, across a range of well-powered studies of different inflectional paradigms, in which children are producing sufficiently high rates of errors to test the prediction, would therefore constitute evidence against this account. German past participles would provide a suitable testing ground. They are formed by up to three markers: the unstressed prefix *ge-;* the weak or strong ending *-t* or *-en,* and, in case of strong verbs, a wide range of changes to the stem vowel. Weak verbs like *machen – gemacht* 'made – made' are productive in contemporary German and have a high type frequency. All strong verbs take the ending *-en,* but there are 15 different types of vowel gradation with 4 to 21 members (count based on personal coding of German CHILDES-corpora), e.g., *kriechen – gekrochen* (crawl – crawled) or *reissen – gerissen* (tear – torn). Previous research has shown that children make errors on the prefix, the suffix, and vowel gradation (Szagun, 2011). Strong verbs have lower type, but high token frequency. Within strong verbs, the suffix *-en* has higher type frequency than the less predictable and more variable vowel gradation. In frequency controlled nonce-verb tasks it should be easier to induce *-en* errors than stem vowel errors.

Productivity through (frequent) morphological frames

Lexical items often occur in morphological frames like *The NOUN VERB-s*, where the slots can be filled by different items. Frequent frames contribute to category formation (Mintz, 2003). In a crosslinguistic study on seven typologically diverse languages Moran et al. (2018) looked at trigrams (A-B-C) and investigated the predictive power categorizing the B-element based on A and C. Irrespective of the language type, the results were most stable for inferring lexical categories like nouns and verbs, presumably because the stable context allowed the children to abstract away from the details of the inflectional form.

Prediction regarding category induction

Morphological frames promote category induction by opening a "similarity space" for the slot-filling element(s). Thus, when controlling for other factors such as semantic and discourse properties – morphological frames with higher transitional frequency (e.g., He's VERB-ing) will be acquired earlier than similar frames with lower transitional frequency (e.g., They're VERB-ing) such that, for example, children will be able to use a novel verb at a younger age in the former frame than the latter in an experimental elicited production study. Repeated failure to find such a difference in suitably controlled and well-powered studies across a range of domains would therefore constitute evidence against the present account.

In the next sections (B, C), I will focus on the acquisition of word order in three different scenarios: fixed word order as in English, variable word order as in German and Dutch. In both cases, the surface patterns are high-frequency schemas, but the underlying form-function relations require a more detailed analysis, based on the integration of different cues. I will argue that both fixed and variable word order languages can be seen as compilations of different constructions that have their own developmental trajectories before network associations are formed. As a third example, word order in complex *wh*-questions is briefly discussed (D). Here, there is limited evidence in the input for the concrete structures, such that children will have to rely on inferences based on simpler component structures.

B. Acquisition of fixed word order languages

Constructivist accounts predict that children will exploit distributional cues to establish form-function mappings. According to the competition model (Bates & MacWhinney, 1987), languages differ as to how many different cues there are to a certain function, how reliable these cues are, and how difficult they are to process in terms of perceptual salience and frequency. The more fixed and canonical word order patterns are, the more predictive are the transitional probabilities (Ibbotson, Solnikov & Walker, 2020). English sentences are highly templatic given the relatively small number of determiners and pronouns, mostly without case distinctions, and verbal inflections. This should facilitate the abstraction of typical word order patterns. Yet even in English learning word order is demanding because variation in word order encodes different functions. Two issues are relevant here: The identification of different sentence types, and the association between position and thematic roles that leads to different interpretations of the sentences.

Word order and sentence type

In an analysis of input addressed to English-speaking children, Cameron-Faulkner, Lieven, and Tomasello (2003) found that only 15% of the input sentence addressed to 2-3-yr-olds had basic SVO word order, whereas the majority of utterances were questions, imperatives, demonstrative copulaic constructions (*that's, there's, here's X)*, or sentence fragments. However, there was a limited number of different sentence onsets with high token-frequency, which facilitates the identification of sentence types.

   At first sight, the simple English transitive poses a problem to the tenet of Construction Grammar that constructions are form-function units, since it seems to be impossible to abstract a meaning for quite different relations. The problem of such "meaningless constructions" is resolved when one assumes that the transitive surface pattern is rather a "hyperschema" that subsumes different form-function mappings or *linking relations* between different verbs and their arguments (cf. Hilpert, 2019, p. 70ff.; examples from Ambridge, 2019, p. 20):

(1) Linking-relations in the English transitive
     Contact (non-causative) [AGENT] [ACTION] [PATIENT]   John hit Bill
     Causative   [CAUSER] [ACTION] [UNDERGOER] John broke the plate
     Experiencer-Theme   [EXPERIENCER] [EXPERIENCE] [THEME] John feared Bill
     Theme-Experiencer   [THEME] [EXPERIENCE] [EXPERIENCER] John frightened Bil
     'Weigh' Construction   [THING] [MEASURE/COST/WEIGH] [AMOUNT] John weighed 100lbs
     'Contain' Construction [CONTAINER] [CONTAIN] [CONTENTS] The tent sleeps four people

Ambridge (2019) follows Abbot-Smith and Tomasello (2006) in arguing that children have to proceed in an exemplar-based fashion and learn these constructions in an item-specific way. Fixed generalizations by, for example, equating the grammatical subject with the agent-role would lead to a wrong interpretation of *John feared Bill*. Only experience can teach children which form-function mappings generalize because they have a prototypical core, and which represent "constructional isolates" or specific pattern that are instantiated by few verbs only.

Sentence interpretation: Association between word order and thematic roles

According to the competition model (Bates & MacWhinney, 1987), languages differ as to how many different cues there are to function, how reliable these cues are, and how

difficult they are to process in terms of perceptual salience. In a cross-linguistic sentence-interpretation task with transitive sentences such as *The dog the cat chases* or The *eraser chases the pig*, Bates et al. (1984) found that depending on the language, speakers weigh cues differently. In English, word order was the most reliable cue for the interpretation of such prototypical transitive sentences for adults. Since children learning English do not hear variable word order in declaratives, and learn to use word order as a strong cue for sentence processing, they do not automatically generalize that a verb heard intransitively (*The cow is meeking*) can be used transitively even if it depicts a transitive action (Brooks & Tomasello, 1999; Tomasello & Brooks, 1998). Children's reliance on item-specific vs verb-general word order was also tested in weird-word-order experiments: When English-speaking children were taught a new verb describing a transitive action in non-canonical word order (SOV: *Elmo the car dacked*; VSO*: Dacked Elmo the car)*, they did not reliably correct it to default SVO-word order until 4 years of age (Akhtar, 1999). The younger groups (2 and 3 year-olds) predominantly copied the word order they heard the particular verb in, but did correct "weird word orders" with familiar verbs like *push* to a higher extent.

Goldberg and Casenhiser (2005) tested when and whether children are ready to accept non-canonical word order if it represents a new function, in this case appearance (*The rabbit the hat mopoed*), where the second NP denotes the location of the appearance (video: rabbit appeared on the hat), and all nonce verbs ended on *-o* to provide a morphological cue (*moopo, vako, suto, keebo, fego*). They found that 5 to 7-yr-old children were able to learn that new construction, in particular when the frequency distribution was skewed towards one verb. Experiments like this show that children become better at fast mapping argument structure constructions with age, especially when morphological and distributional cues support unambiguous generalization.

Prediction

The acquisition of fixed word order is facilitated by the templatic nature of the input patterns. However, even simple transitive sentences represents a number of different constructions in terms of form-function mappings, such that too early generalizations based on highly frequent or prototypical linking patterns could hinder the acquisition of non-prototypical and/or low frequency patterns. Whether learners at different stages treat different thematic role assignments (e.g., experiencer-theme vs theme-experience) with the same surface word order (SVO) as the same or as different constructions could be tested by sentence processing experiments. In the visual-world paradigm, eye-tracking information is used to trace the online processing. When analyzing sentences like *Susan frightened John* vs *Susan feared John*, the method could be used to see if children have an "agent-first" interpretation strategy, or whether they wait for the verb to identify the correct thematic role.

C. Acquisition of "movement constructions" in variable word order languages

In Construction Grammar no movement operations are assumed. Instead, linguistic units that differ in word order *and* function constitute different constructions with their respective functions. The basic hypothesis is again that each construction would be acquired from exemplars of that same construction. In principle, these constructional schemas could be independent of one another, but since they overlap in lexical material

and co-occur in conversations, children should be able to infer the relationship between different constructions based on distributional evidence (see Diessel, 2019, for a survey of types of network relations). The acquisition problem addressed in this section deals specifically with the different kinds of network relations that result from stable form-function and position-mappings in languages that show a more variable word order.

Küntay and Slobin (1996, p. 267) identified **variation sets** in the input where the communicative goal is kept constant but the form varies, as in these examples from a father to his 2;3 year-old son:

(2) *Who did we see when we went to the store?*
Who did we see?
Who did we see in the store?
Who did we see today?
When we went out to shopping, who did we see?

Such variation sets help children to build paradigmatic and syntagmatic relations in that they realize which elements can replace one another, and also how structures can be expanded or moved around, in this case for pragmatic topicalization. At the same time, the stable core of the utterance schema is entrenched through repetition.

Similar variation occurs in languages that have variable word order. A classic case is the so-called verb-second phenomenon in some West-Germanic languages like Dutch and German. Finite and nonfinite parts of the inflected verbs occur in different positions and form a so-called "sentential bracket" or *Satzklammer* (3). This leads to five topological fields: Finite verbs occur in second position (declaratives, 3.a and b) or sentence initially in yes/no-questions (3.c), but in final position in subclauses (d). When the verb is complex (either a separable particle verb or a main verb plus auxiliary), the verb complex is separated by the middle (3 b,c). With few exceptions, only one constituent occupies the pre-field. The pre-field can be occupied by the subject and wh-pronouns like in English, but also by other topicalized constituents (3.b). As a consequence, most arguments and adjuncts fill the middle-field, which leads to discontinuity between the finite and non-finite part of complex verbs, though extraposition to the post-field is possible and reduces discontinuity (3.b).

(3) The topological model of sentence structure in German

pre-field	V1/V2	middle field	V-final	post-field

a. **Simplex verb: SVO in main clauses**

| Peter | | bringt dem Kind Geschenke | | |
| Peter | | brings the child gifts | | |

b. **Sentential bracket with complex verbs or auxiliary + verb**

| Peter | | bringt dem Kind Geschenke | mit | |
| Peter | | brings the child gifts | along | |

| Dem Kind | will | Peter Geschenke | bringen | zum Spielen |
| The child-DAT | will | Peter-NOM gifts | bring | to play |

| Geschenke | hat | Peter dem Kind | mitgebracht | |
| Gifts-ACC | has | Peter the child | along-brought | |

c. *yes/no-questions*:

Bringt Peter dem Kind Geschenke mit?
Brings Peter the   child gifts with?

Wird   Peter dem Kind Geschenke bringen?
Will   Peter the child gifts bring?

d. *subclause: SOV*

…, weil   er dem Kind Geschenke bringen will.
…,   because he the child gifts   bring will.

Learners have to deal with the discontinuous nature of complex verbs and to identify the regular association between type of inflection (finite/non-finite) and position. This phenomenon has attracted considerable attention in the generative literature where it is assumed that the finite verb raises from the finite SOV-base position to the V1/V2 position in main clauses ("verb second effect"). As this phenomenon is attested in some (German, Dutch, Afrikaans), but not all Germanic languages, it was argued that the type and extent of verb movement is licensed by aspects of verb morphology (Ambridge & Lieven, 2011; Blom, 2007; Lasser, 2002; Westergaard, 2009).

Alternatively, the sentential bracket can be described as a coalition of different sentence structures, just like the hyperschema for English transitives. The basic distributional facts are quite simple: With few exceptions, only one constituent can, but need not precede the finite verb in declarative sentences and questions, and the non-finite verb or verb particle is typically found at the end of the clause. Traditional grammar accounts identified 28 basic syntactic templates, based on argument structure and the distributional constraints of the verbs (Wöllstein-Leisten, Hilmann, Stephan & Vikner, 1997). There is ample evidence for the position of finite and non-finite verbs in the input which should allow for rapid pattern abstraction. Behrens (2006) computed the production rate in a dense-database of child and child-directed speech, when the child was between 4;4 and 4;11. In a lively conversation, the child hears and produces several hundred verb forms per hour (4).

(4) Production rate *per hour* (based on a 40-hour sample)

	Child	Adults
Main verbs	209	469
Copulas	58	120
Auxiliaries	49	136

However, like in English, the acquisition problem becomes more complex when semantic-pragmatic effects of topicalization and extraposition are concerned (see below). From a Construction Grammar perspective, verb-first-constructions like yes-/no-questions, imperatives, and certain types of conditionals can be characterized by their pragmatic properties as non-assertive (Diessel, 1997). I.e., in adult German we see word order variation that is linked to sentence type (question, declarative, imperative; main clause vs. subclause) and morphology that encodes semantic and pragmatic distinctions (agreement, tense, modality). So rather than seeing word order variation as a uniform phenomenon where the verb "moves" depending on sentence type and verb inflection, one can conceive of German word order as a family of constructions where word order variation encodes different types of information, but

where these constructions converge on the limitations imposed by the schematic nature of the sentential bracket. Hence, Construction Grammar does not assume that there is a basic word order from which others are derived by movement, but rather sees different word order patterns as schematic representations that encode different sentence types as well as information structure (Boas & Ziem, 2018; Welke, 2019). This entails that the learning problem turns from the acquisition of a rule that licenses verb-movement to one where numerous constructional schemas have to be acquired (i.e., the form-function correspondences between different sentence types and inflectional markers).

In the acquisition of word order in German and Dutch we see a remarkable developmental change in children's grammar in the second year of life: children initially predominantly use non-finite particles, infinitives and past participles in sentence-final word order, and a smaller set of finite verbs in correct V1/2-position (Behrens, 1993; Poeppel & Wexler, 1993; Wijnen, Kempen & Gillis, 2001). Previous studies identified several sources of information in the input that contribute to building syntagmatic and paradigmatic relations, but their exact interplay across developmental time has yet to be determined. Based on these findings, I will sketch a scenario with changing predictors over developmental time, where finite and non-finite verbs and their arguments will be acquired as different clusters, and where the frequency of specific verb forms and their left and right neighbors interacts with an attentional focus on the beginning or end of the utterance and a growing processing window.

Sentence processing: Learning from the right and left periphery

Attention tuning, to find out where and how the relevant information is encoded, is essential for language learning (Ellis, 2006a, 2006b; Slobin, 1985). Utterance-initial and utterance-final positions are particularly salient, such that tuning in on these positions provides the child with strong cues about word order. This was shown in a computational modelling study of the acquisition of word order where the model expanded the processing window with either a focus on the left periphery for English and Spanish, or a focus on the right periphery for Dutch and German (Freudenthal, Pine, Aguado-Orea & Gobet, 2007).

Distributional evidence from utterance-final clusters with non-finite verb forms

A large proportion of early verb forms in German and Dutch are nonfinite infinitives or past participles (*runtergefallen* 'down-fallen'), often preceded by the topic object (*Ball spielen* 'ball play-INF'). The predominance of infinitives in early child Dutch or German is an effect of their salient sentence final position and the semantically heavy content of lexical verbs (activity, achievement and accomplishment verbs). Non-finite verb forms are typically preceded by direct and indirect objects, hence the end of the clause is occupied by lexical elements rich in information about the event itself. In the absence of finiteness markers for agreement, tense and modality, most utterances have what researchers called a "modal" or non-actual reading because it is not anchored on the time-line: An utterance like *Eis essen* 'ice cream eat-INF' encodes deontic modality as in *I want to eat ice cream* (Blom, 2007; Hoekstra & Hyams, 1998; |Jordens, 1988, 2002; Lasser, 1997; Schimke & Dimroth, 2018).

In sum, children derive clusters rich in information from the right periphery of German clauses that encode the "comment" in terms of topic-comment information structure. They will produce truncated subject- or object-verb clusters with activity and telic verbs that are frequently attested in clause-final position. By now, this is quite a robust finding in the acquisition literature (Freudenthal, Pine & Gobet, 2009; Wijnen et al., 2001). A new take on this distributional facts is to study the development over time, in particular how children become flexible with using the same lexical items in both positions (see below). In order to develop this argument, we will first look at how children acquire the left periphery of German clauses where finite verbs are found.

Distributional evidence from utterance-initial clusters with finite verb forms

In contrast to the event-centered information at the right periphery, utterance initial clusters provide information on the semantic status of the utterance (sentence type, tense and modality, and negation, among others). There is evidence that early on, finite verb forms constitute a smaller and functionally different set of verbs: In Dutch, early finite verbs are mostly (modal) auxiliaries or static verbs like the copula *be* and the verb *to have* (de Haan, 1987). Schimke and Dimroth (2018) found similar effects of verb semantics in German: semantically light like *make, do, come, go* verbs occur with finiteness marking, full lexical verbs in non-finite forms. This corresponds to recent findings from spoken adult corpora in German, Dutch, and English, where finite forms are much more frequent than nonfinite forms, but also lexically restricted to a small set of high-frequency, but semantically light verbs or auxiliaries. Kempen and Harbusch (2019) argue that the predominance of only a few high-frequency verbs early in the utterance facilitates processing. When looking at their list of the 50 most frequent German finite verbs in written corpora, it seems that many of them are not typical in CDS such that the set of high-frequency finite verbs addressed to children should be even smaller. This would also substantially reduce the problem of learning verbal inflection because children only have to deal with a small set of verbs that carry agreement and tense marking. A further facilitating factor could be the frequent use of separable particle verbs like *reinmachen – mach X rein* 'put in' (Behrens, 1998). They contribute massively to the lexical richness, but are based on a smaller number of high-frequency and often semantically light verb roots, which adds further support to the hypothesis that the set of finite verbs is more limited than the set of nonfinite forms found clause-finally.

Most likely, we should see effects of specific inflected word forms, not of the verb lemma in general. In German, many high frequency verbs are semantically light and polysemous. Finite and and non-finite forms not only occur in different positions, but also with different functions (*have* or *go* as auxiliary or main verb) or semantics / argument structure. There is evidence that item-effects occur at the level of word forms, not across all different usage-types involving the same verb lemma (Behrens, 2003; Lange, 2007; Theakston, Lieven, Pine & Rowland, 2002).

In terms of information structure, the onset of utterances provides evidence of the sentence type (question, declarative or imperative), and hosts negation (*nicht* 'not') as well as additive particles like *auch* 'too, also' used to be included in activities (*ich will auch spielen* 'I want to play, too'). These particles could pave the way for finite verb forms because the typically occur after the finite verb (*das geht nicht/auch rein* 'that goes not/also in'), and encode relevant information regarding the intention of the

speaker (negation, addition/inclusion). This allows learners to recognize that relevant types of information regarding speech act type, negation or addition/inclusion are encoded clause-initially. Schimke and Dimroth (2018) show for that first language learners, these particles operate like a boundary or pivot when they realize that the semantically light and phonologically non-salient finite verb precedes these particles. Hence, finite verbs form local clusters with preceding onset subject (pro)nouns and/ or adjacent focus/negation particles (5).

(5) Prominent schemas in the left sentence periphery

(Pro)noun	V-FIN		(*nicht/auch*)	declaratives
	V-FIN	(pro)noun	(*nicht/auch*)	y/n-questions
Wh-word	V-FIN		(*nicht/auch*)	wh-questions

Prediction for the acquisition of finite verbforms at the left periphery

Finite verb forms are acquired based on a limited number of local, clause initial clusters. When tested for their knowledge of agreement and tense inflection, children should make fewer errors with high-frequency verb forms attested in this position, than with low-frequency ones. A failure to find such token-frequency effects, but verb-general success in agreement marking would provide counterevidence.

In sum, early German and Dutch child language shows a strong correlation between verb form and position, and an effect of complexity reduction in CDS: There are significant associations between verb position and verb semantics. Modal, auxiliary, and static verbs tend to occur as finite verbs in V1/V2 position, whereas activity and telic verbs tend to occur as infinitives and participles at the high-saliency end of utterances (Wijnen et al., 2001). Thus, learners are not confronted with all possible variations in the input, but with a limited set of lexical frames, which facilitates pattern abstraction. Also, children find different types of information at the onset and end of utterances, and this tunes their attention to both the end and the beginning of utterances, and should allow them to make inferences about the pragmatics of information structure. In a series of computational modellings, the acquisition of sentence structures in German, Dutch, Spanish and English could be simulated by installing a limited working memory with an attentional focus to the beginning and/or end of utterances to model primacy and recency effects (Freudenthal, Pine, Jones, and Gobet (2015). They explain the emerging word order patterns by attention tuning and extending the working memory of the model, which allows the processing of increasingly larger clusters. In the following, I will spell out how this might work for German and Dutch word order.

Becoming productive: Effects of semantics and lexical overlap

Despite the distributional biases discussed above, *all* German and Dutch verbs can occur in finite or non-finite forms, hence children will have to expand the range of verbs attested in different constructions. And they will have to learn to produce longer sentences with finite AND non-finite verbs.

Initially, we do not expect much overlap between verbs occurring in finite form and position and those occurring in non-finite form and position. Then children learn more constructional schemas, and the fact that the same verb can occur in either position, with the appropriate morphological finiteness-marking. Whether and how children

build productive associations between the different verb positions should depend on the amount of overlap of the same verb in different tenses, for example present perfect with discontinuous verb position and simple past with fronted position as in (6):

(6) *Hast du mir ein Geschenk gebracht?*   'Did you bring me a gift?'
    Bringst du mir ein Geschenk?   'Will you bring me a gift?'

As in the variation sets described above, the syntactic relationship between the two word order patterns should be clearer when the function is similar. This is particularly clear with separable particle verbs where the semantics is the same in continuous or discontinuous order (7):

(7) *Was machst du da rein?*   *Was hast du da reingemacht?*
    'what do you put in?'   'what have you there in-put?'

However, many verbs are highly polysemous, and this can obscure the formal relationships between early utterance schemas such that children proceed with not just lexeme-specific frames, but even wordform-specific frames. The verb *to go*, for example, is used in different functions that are associated with different inflectional forms (*going, gone, goes, went* etc.). Distributional analyses suggest that children do not link these different schemas, as would be indicated by flexible variation of word order and inflection, but use them as wordform-specific templates (Behrens, 2003; Theakston et al., 2002). Both studies found that different forms of the verb *go* in English, Dutch and German appear in their own distributional environments with different functions.

Prediction

Learning the association between verb form and position can facilitated by overlap in form, but inhibited by differences in semantics, even at the level of inflectional forms. Separable particle verbs could be path-breaking elements for lexical overlap, as their semantics is stable, and as they are formed with a limited number of high-frequency verb roots. We expect the first verbs to be used in finite and nonfinite forms to be high-frequency, and within the same semantic function. To find a large set of semantically different verbs used in both positions at an early stage of development would provide counterevidence.

Accounting for Agreement Errors with Transitional Probabilities

Transitional probabilities help to explain both correct and incorrect patterns. Poeppel and Wexler (1993), in a case study of a single transcript of a boy aged 2;1, argued that even at this early age, there were hardly any agreement errors between subject and verb. They consider this as evidence that the finite inflection of the verb licenses subjects. From a constructivist position, the argument could be reversed: A given subject could trigger the corresponding verb form because children build up verb frames (NAME/pronoun V-AGR). If this is the case, we expect lower error rates in well-established units with an overt subject than in clusters where the subject is elided or infrequent.

Subsequent studies reported agreement errors in German and Dutch where a subject is used with a non-finite verb (so-called Optional Infinitive or OI-errors) as in English *he go*

or German *Mama machen* 'mummy make-INF'. Whereas the frequency of such OI errors does not depend on the sentence type (declarative or wh-question) in English (OI errors frequent in both constructions) or Spanish (OI errors *in*frequent in both constructions), they do in Dutch and German: Here, OI-errors are infrequent with WH-questions, and more frequent in declaratives (Freudenthal et al., 2015). This can be explained by the distributional facts imposed by the topological model: In German and Dutch, a wh-element is *always* followed by a finite verb. This leads to the prediction that children can form frames like *wh-word …s* for 3rd person singular, based on transitional probabilities. The only source of confusion is that verbs in present tense plural (1st and 3rd person) are homophonous with the infinitive (cf. *Wohin gehen sie*) 'where go-3p they?' and *Wohin sollen sie gehen*? 'where shall-INF they go-INF')

The picture sketched here is that learning "verb-second" in German or Dutch is not a uniform process where inflectional properties license verb movement, as proposed in the generative literature from the 1990ies. Instead, the topological model of sentence structure serves as a stable hyperschema that specifies the word order of different sentence types. Acquisition starts from utterance-initial and -final positions, as evidenced by acquisitional sequences and error patterns that can be predicted from local clusters and their transitional probabilities. The existing literature shows that the development from a mainly non-finite to a finite system takes place in the 3rd year of life, based on abundant evidence sentence in the input. But how does a constructivist account fare for low-frequency target structures?

D. Acquisition of complex structures with limited evidence in the input

The "logical problem" of language acquisition does not exist in a usage-based perspective (Hendriks, 2000; Pullum & Scholz, 2002). If there are no innate constraints on possible syntactic structures, there cannot be a mismatch between what the child knows and what s/he can derive from the input. This does not imply that children need to hear all possible sentences or all possible structures of a language before they can process them. Rather, gaps in the evidence must be filled by inferencing, i.e., going beyond the input and generalize over (a combination of) functional and formal aspects relevant for the construction at stake, or INDIRECT POSITIVE EVIDENCE (Lewis & Elman, 2001b).

Constructivist approaches propose that learners can induce such complex structures deriving knowledge about grammatical and ungrammatical sequences from transitional probabilities (e.g., Reali & Christiansen, 2005), and by schema combinations, i.e., building up the respective knowledge from simpler, related structures (Abbot-Smith & Behrens, 2006; Ambridge, Rowland & Gummery, 2020; Dąbrowska, 2008, 2015; Lewis & Elman, 2001a). I will illustrate the sources children can draw on with the acquisition of complex yes-/no-questions in English.

Complex yes-/no-questions

When asking a yes-/no-question with an embedded relative clause, subject-verb inversion is required (8 a,b). In generative grammar, this is a structure dependent operation that respects the syntactic hierarchy by fronting the auxiliary of the main clause, leaving the embedded clause (underlined) unaffected. Errors as in (8.c) supposedly can only be prevented if the speaker is aware of the structure dependence.

(8) a. *The boy <u>who is smoking</u> is crazy.*
    b. *Is the boy <u>who is smoking</u> crazy?*
    c. *Is the boy <u>who smoking</u> is crazy?*

The hypothesis that children should commit errors as in (8.c) is based on the assumption that learners process the structure in terms of abstract categories (i.e., processes the two occurrences of *is* as AUX), and apply a movement rule, but erroneously move the first auxiliary if they are not aware of the syntactic embeddedness of the subject relative. None of these assumptions is made by constructivism. Instead, a much simpler analysis of the necessary operation is provided: The familiar NP *the boy* is elaborated with a relative clause, and this forms the syntactic and semantic unit *The boy who is smoking*. The syntactic process required to produce (8.b) boils down to slotting the complex NP into the open slot of the *yes/no*-question construction that already has the correct word order (*Is X crazy?*), and does not require movement (Ambridge et al., 2020; Ambridge, Rowland & Pine, 2008; Dąbrowska, 2008, 2010). In sum, the knowledge of the *yes/no*-question construction plus the knowledge of relative clause structures should suffice to process these sentences even if their combination is very rare or non-existent in the input.

Transitional probabilities (n-gram statistics).

Reali and Christiansen (2005) provide a more elaborate discussion of what can count as "relevant data" for the acquisition of such constructions, and provide statistical evidence from actual input data from the CHILDES corpus. They used simple bigram- and trigram-statistics (two-word and three-word sequences) for transitional probabilities between words and as evidence for chunks. Note that, in (8.c), the incorrect auxiliary extraction would lead to the non-attested string *who smoking*. Statistical analyses of the input and the child data showed that statistical learning based on bigrams and trigrams successfully predicted grammatical and ungrammatical auxiliary extraction. In a further network simulation using the same data, Reali and Christiansen (2005) showed how inferences drawn from simple transitional probabilities could additionally lead to part-of-speech-category induction by means of multiple cue integration. However, subsequent research showed that the statistical predictions derived from bigrams alone are not as reliable when other inversion structures, or structures involving a different or wider range of verbs and tenses were analyzed. In addition, the results – typically based on small corpora – differed considerably depending on which corpus was used and what analyses were carried out, raising the question of what is a suitable reference corpus for input studies (see Fitz & Chang, (2017) for a review of the critical issues, and additional cues by the inclusion of semantic relations). While transitional probabilities alone might not suffice to account for complex constructions, schema combination-accounts promise to be more successful, because they also preserve the knowledge about the form-function correspondences of the component structures.

Schema combination (Induction from precursor structures).

Lewis and Elman (2001b) trained a Simple Recurrent Network to infer such complex yes/no-questions as in (8) from simpler constructions that provided the system with

all necessary grammatical information (agreement, yes/no-questions in simple sentences, complex sentences with embedded relative clauses).

In a similar vein, Ambridge, et al. (2020) conducted a training study to test whether children can learn the structure in (9.c) when trained on the component schemas (9.a and 9.b), which are attested in the input to children:

(9) a. Is [THING] [ACTION]ing?   e.g., *Is the bird cleaning?*
    b. The [THING] who's [PROPERTY]   e.g., *The bird who is sad.*
    c. Is [the [THING] who's [PROPERTY]] ACTIONing? e.g., *Is the bird who is*
                                                    sad cleaning?

The study with 4- to 6-year-old children was successful in that children in the experimental group where significantly better in producing structures with the trained relative pronoun *who*. However, children in both the experimental and the control group (which was trained with simple NPs like *the sag bird* instead of (9.b) frequently produced relative clauses with *that,* the more colloquial option (*the bird that is sad*). When these were included as correct instances of the target structure, the advantage of the experimental group disappeared as even the children in the control group produced complex questions with embeddings. This suggests a contribution of lexical-specificity and structural priming: In this age range, most children are able to produce questions with embedded clauses with *that*, and – upon training – also with lower-frequency *who*.

Prediction

Complex constructions with low input frequency can be acquired on the basis of indirect positive evidence derived from their component parts. Transitional probabilities provide evidence for likely or unlikely combinations, and schema combination simplifies the task. For example, Abbot-Smith and Behrens (2006) found different acquisition rates for German *werden-* and *sein*-passives (*werden/sein* + past participle). Both are equally complex and frequent in the input, but the *sein*-passive profited from the previous acquisition of its component parts (past participles and the verb *sein* used as copula and auxiliary). They also found that future constructions (auxiliary *werden* + infinitive) were learned later than *werden*-passives despite similar input frequency. Here, one would hypothesize to find fewer inflectional and lexical constraints than in the earlier-learned constructions, because children should be able to transfer their knowledge from already acquired related construction. To find productive complex constructions in child language without assessing the simpler constructions first would provide counterevidence for this constructivist account.

Back to theory: The nature of linguistic processing and representation: shallow or hierarchical?

Constructivist approaches question the underlying motivations of structuralist and generativist linguistics, that linguistic generalizations have to be as powerful as possible. If children learn from local clusters, and manage to generalize on the basis of quite concrete form-function associations, are "abstract" generalizations needed and attested at all, and what is their added value (see Langacker's rule-list fallacy

above)? In other words: Are the generalizations that linguists derive the same as those made by speakers, and is processing always based on the most abstract available representation?

Exemplar-storage and low-level analogies

Within usage-based approaches, the general idea is that initial generalizations are "concrete" or item-based (i.e., chunks or *n*-grams) and turn more "abstract" later (from slot & frame-patterns with some variability to more productive semantically-based constructions (agent/patient) and finally to rules based on grammatical categories (subject, object, accusative, dative etc.). Ambridge (2019) criticizes usage-based accounts for accepting the view that the endstate of language development are abstract rules, and the implication that the nature of processing changes: after certain (to be defined) productivity thresholds, the mind would start to translate the words into more abstract categories – from *kisser* to *agent* to *subject*. But processing studies with children and adults, as well as computational models of early language acquisition suggest that processing based on concrete units of language (words and n-grams or multiword strings/chunks) is successful, and often superior to rule-based accounts. Such exemplar-effects last into adulthood. In a literature review, McCauley and Christiansen (2019) show that entrenched multiword units are processed faster in comprehension and production. They argue that multiword units "retain their significance throughout development" (McCauley & Christiansen, 2019, p. 3). So-called "shallow" or non-hierarchical, underspecified processing in adults such suggests that there is more continuity between child and adult processing than previously assumed (McCauley & Christiansen, 2019, p. 4).

In a similar vein, Ambridge (2019) argues in favor of radical exemplar-based processing because the minimal variation of stored exemplars based on analogy bypasses a number of problems of accounts that assume processing by rule. If low-level generalization were stored schemas, they could be retrieved and used with minimal adaptations in similar communicative contexts. The basic argument ("lumping and splitting") is that the more general an abstraction, the more heterogeneous cases it subsumes and the lower its predictiveness (see the low predictiveness / high heterogeneity of hyperschemas discussed above) such that the general categories would have to be re-represented as different subtypes. The radical versions of exemplar- and usage-based model seem controversial because the essence of learning is making predictions based on generalizations (Ambridge, 2020). But it makes developmental sense if one starts out with the assumption that the child has to build-up experience first. Exemplar-based abstraction leads to precise generalizations based on similarity because one has to learn the details before one can lump. At start, it does not matter whether a certain construction is generalizable or not because the child does not have a baseline for comparisons yet.

Linguistic theory

Exemplar-processing has informed constructivist linguistic theory regarding language typology, and historical and developmental language change. Croft (2001) and Diessel (2019) argue that there are no universal categories, and that the scope of possible generalizations is typically quite limited if one looks at the totality of constructions in a language. Diessel (2019) integrates findings from language change

and processing to identify the forces that lead to grammatical relations. He argues that linguistic relation can be accounted for in terms of symbolic relations (form-function associations), sequential relations (that also lead to categorization or unit-formation), and taxonomic relations (hierarchical inheritance links between constructions). Taxonomic relations capture the similarities and specifics between different genitive constructions, for example.

  Schmid's entrenchment and conventionalization model (Schmid, 2015, 2020) adds a social dimension that promotes and constrains generalization. While entrenchment happens in the individual mind, conventionalization requires convergent generalizations between speakers. Currently, there is little research on the social mechanisms that drive language learning, be it in terms of inferencing the interlocutor's intentions (Matthews, 2014) or in making use of discourse-level information (Casillas & Frank, 2017). Many arguments for constructivist accounts are based on large databases of naturalistic interactions. A closer look at the concrete usage-situations will help us to refine the effect of distributional information by looking at the pragmatic and functional properties in more detail.

References

Abbot-Smith, K., & Behrens, H. (2006). How known construction influence the acquisition of other constructions: the German periphrastic passive and future constructions. *Cognitive Science*, *30*, 995–1026.

Abbot-Smith, K., & Tomasello, M. (2006). Exemplar versus prototype models in syntactic acquisition. *The Linguistic Review*, *23*, 275–290.

Akhtar, N. (1999). Acquiring basic word order: Evidence for data-driven learning of syntactic structure. *Journal of Child Language*, *26*(2), 339–356.

Allen, S. (2017). Polysynthesis in the acquisition of Inuit languages. In M. Fortescue, M. Mithun, & N. Evans (Eds.), *Handbook of Polysynthesis* (pp. 449–472). Oxford: Oxford University Press.

Allen, S., & Behrens, H. (2019). Insights into understanding human language from children's acquisition of morphology and syntax: A historical and current perspective on central questions in the field. In P. Hagoort (Ed.), *Human Language: From Genes and Brains to Behavior* (pp. 127–145). Cambridge, MA: MIT Press.

Ambridge, B. (2019). Against stored abstractions: A radical exemplar model of language acquisition. *First Language.* doi:10.1177/0142723719869731

Ambridge, B. (2020). Abstractions made of exemplars or 'You're all right, and I've changed my mind': Response to commentators. *First Language*, *40*(5–6), 640–659. doi:10.1177/0142723720949723

Ambridge, B., & Lieven, E. (2011). *Child language acquisition: Contrasting theoretical approaches.* Cambridge: Cambridge University Press.

Ambridge, B., Rowland, C. F., & Gummery, A. (2020). Teaching the unlearnable: A training study of complex yes/no questions. *Language and Cognition*, *12*(2), 385–410. doi:10.1017/langcog.2020.5

Ambridge, B., Rowland, C. F., & Pine, J. M. (2008). Is structure dependence an innate constraint? New experimental evidence from children's complex-question production. *Cognitive Science*, *32*, 222–257.

Bates, E., & MacWhinney, B. (1987). Competition, variation and language learning. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 157–193). Hillsdale, NJ: Erlbaum.

Bates, E., MacWhinney, B., Caselli, C., Devescovi, A., Natale, F., & Venza, V. (1984). A crosslinguistic study of the development of sentence interpretation strategies. *Child Development*, *55*, 341–354.

Behrens, H. (1993). *Temporal reference in German child language: form and function of early verb use.* Universiteit van Amsterdam: Doctoral dissertation, Instituut voor Algemene Taalwetenschap.

Behrens, H. (1998). How difficult are complex verbs: evidence from German, English and Dutch. *Linguistics (Special Issue: Developing a verb category: crosslinguistic perspectives)*, *36*, 679–712.

Behrens, H. (2003). Bedeutungserwerb, Grammatikalisierung und Polysemie: Zum Erwerb von "gehen" im Deutschen, Niederländischen und Englischen. In S. Haberzettl & H. Wegener (Eds.), *Spracherwerb und Konzeptualisierung* (pp. 161–181). Frankfurt/M.: Peter Lang.

Behrens, H. (2006). The input-output relationship in first language acquisition. *Language and Cognitive Processes*, 21(1–3), 2–24. doi:10.1080/01690960400001721

Blom, E. (2007). *From Root Infinitive to Finite Sentence: The Acquisition of Verbal Inflections and Auxiliaries*. Berlin/New York: Mouton de Gruyter.

Boas, H. C., & Ziem, A. (2018). Approaching German syntax from a constructionist perspective. In H. C. Boas & A. Ziem (Eds.), *Constructional Approaches to Syntactic Structures in German* (pp. 181–203). Berlin/Boston: De Gruyter

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

Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.

Bruner, J. (1983). *Children's talk: learning to use language*. New York: Norton & Company.

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

Cameron-Faulkner, T., Lieven, E., & Tomasello, M. (2003). A construction based analysis of child directed speech. *Cognitive Science*, 27(6), 843–873. doi:10.1016/J.Cogsci2003.06.001

Casillas, M. & Frank, M. C. (2017). The development of children's ability to track and predict turn structure in conversation. *Journal of Memory and Language*, 92, 234–253. doi:10.1016/j.jml.2016.06.013

Christiansen, M. H., & Chater, N. (2016). *Creating language: Integrating evolution, acquisition, and processing*. Cambridge, MA: MIT Press.

Clark, E. V., & Nikitina, T. N. (2009). One vs. more than one: Antecedents to plurality in early language acquisition. *Linguistics*, 47(1), 103–139. doi:10.1515/LING.2009.004

Croft, W. (2001). *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford: Oxford University Press.

Dąbrowska, E. (2008). Questions with long-distance dependencies: A usage-based perspective. *Cognitive Linguistics*, 19(3). doi:10.1515/cogl.2008.015

Dąbrowska, E. (2010). The mean lean grammar machine meets the human mind: Empirical investigations of the mental status of linguistic rules. In H.-J. Schmid & S. Handl (Eds.), *Cognitive Foundations of Linguistic Usage Patterns: Empirical Studies* (pp. 149–170).

Dąbrowska, E. (2012). Different speakers, different grammars: Individual differences in native language attainment. *Linguistic Approaches to Bilingualism*, 2(3), 219–253. doi:10.1075/lab.2.3.01dab

Dąbrowska, E. (2015). What exactly is Universal Grammar, and has anyone seen it? *Frontiers in Psychology*, 6, 852. doi:10.3389/fpsyg.2015.00852

Dąbrowska, E. (2020). Language as a phenomenon of the third kind. *Cognitive Linguistics*, 31(2), 213–229. doi:doi.org/10.1515/cog-2019-0029

de Haan, G. J. (1987). A theory-bound approach to the acquisition of verb placement. In G. D. Haan & W. Zonneveld (Eds.), *Formal Parameters of Generative Grammar (OTS Yearbook III)* (pp. 15–30). Dordrecht: ICG.

Diessel, H. (1997). Verb-first constructions in German. In M. Verspoor, L. K. Dong, & E. Sweetser (Eds.), *Lexical and Syntactical Constructions and the Construction of Meaning* (pp. 51–68). Amsterdam: Benjamins.

Diessel, H. (2019). *The Grammar Network*. Cambridge: Cambridge University Press.

Ellis, N. C. (2006a). Language acquisition as rational contingency learning. *Applied Linguistics*, 27(1), 1–24. doi:10.1093/applin/ami038

Ellis, N. C. (2006b). Selective attention and transfer phenomena in L2 acquisition: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. *Applied Linguistics*, 27(2), 164–194. doi:doi:10.1093/applin/aml015

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

Fillmore, C. J. (1988). The mechanisms of Construction Grammar. *BLS*, 14, 35–55.

Fitz, H., & Chang, F. (2017). Meaningful questions: The acquisition of auxiliary inversion in a connectionist model of sentence production. *Cognition*, 166, 225–250. doi:10.1016/j.cognition.2017.05.008

Freudenthal, D., Pine, J. M., Aguado-Orea, J., & Gobet, F. (2007). Modelling the developmental patterning of finiteness marking in English, Dutch, German and Spanish using MOSAIC. *Cognitive Science*, 31, 311–341.

Freudenthal, D., Pine, J. M., & Gobet, F. (2009). Simulating the referential properties of Dutch, German, and English Root Infinitives in MOSAIC. *Language learning and development*, 5(1), 1–29. doi:10.1080/15475440802502437

Freudenthal, D., Pine, J. M., Jones, G., & Gobet, F. (2015). Simulating the cross-linguistic pattern of Optional Infinitive errors in children's declaratives and Wh- questions. *Cognition*, *143*, 61–76. doi:10.1016/j.cognition.2015.05.027

Goldberg, A. E. (1995). *Constructions*. Chicago: Chicago University Press.

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

Greenhill, S. J., Wu, C. H., Hua, X., Dunn, M., Levinson, S. C., & Gray, R. D. (2017). Evolutionary dynamics of language systems. *Proc Natl Acad Sci U S A*, *114*(42), E8822–E8829. doi:10.1073/pnas.1700388114

Hendriks, P. (2000). The problem with logic in the logical problem of language acquisition. In L. R. Gleitman & A. Joshi (Eds.), *Proceedings of the 22nd Annual Meeting of the Cognitive Science Society* (pp. 220–225). Hillsdale, NJ: Erlbaum.

Hilpert, M. (2019). *Construction Grammar and its Application to English* (2). Edinburgh: Edinburgh University Press.

Hoekstra, T., & Hyams, N. (1998). Aspects of root infinitives. *Lingua*, *106*, 81–112.

Ibbotson, P. (2013). The scope of usage-based theory. *Frontiers in Psychology*, *4*, 255. doi:10.3389/fpsyg.2013.00255

Ibbotson, P. (2020). *What it Takes to Talk*. Berlin: Mouton de Gruyter.

Jordens, P. (1988). The aquisition of word order in L2 Dutch and German. In P. Jordens & J. Lalleman (Eds.), *Language development*. Dordrecht: Foris.

Jordens, P. (2002). Finiteness in early child Dutch. *Linguistics*, *40*, 687–765.

Kempen, G., & Harbusch, K. (2019). Mutual attraction between high-frequency verbs and clause types with finite verbs in early positions: corpus evidence from spoken English, Dutch, and German. *Language, Cognition and Neuroscience*, *34*(9), 1140–1151. doi:10.1080/23273798.2019.1642498

Kidd, E., & Donnelly, S. (2020). Individual differences in first language acquisition. *Annual Review of Linguistics*, *6*:8.1–8.22. doi:doi.org/annurev-linguistics-011619-030326

Kidd, E., Donnelly, S., & Christiansen, M. H. (2018). Individual differences in language acquisition and processing. *Trends Cogn Sci*, *22*(2), 154–169. doi:10.1016/j.tics.2017.11.006

Kirby, S. (2017). Culture and biology in the origins of linguistic structure. *Psychon Bull Rev*, *24*(1), 118–137. doi:10.3758/s13423-016-1166-7

Kirov, C., & Cotterell, R. (2018). Recurrent neural networks in linguistic theory: Revisiting Pinker and Prince (1988) and the past tense debate. *Transactions of the Association for Computational Linguistics* (6), 661–665.

Küntay, A., & Slobin, D. I. (1996). Listening to a Turkish mother: Some puzzles for acquisition. In D. I. Slobin, J. Gerhardt, J. Guo, & A. Kyratzis (Eds.), *Social interaction, social context, and language: Essays in honor of Susan Ervin-Tripp* (pp. 265–286). Hillsdale NJ: Erlbaum.

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

Langacker, R. W. (2000). A dynamic usage-based model. In M. Barlow & S. Kemmer (Eds.), *Usage-based Models of Language* (pp. 1–63). Stanford: CSLI Publications.

Lange, B. (2007). *machen, haben, gehen, kommen. Einige "Passepartout"-Verben im Primärspracherwerb des Deutschen.* (Vol. *45*). Frankfurt am Main: Peter Lang.

Lasser, I. (1997). The interpretation of root infinitive constructions in adult and child German. In J. C. Schaeffer (Ed.), *The interpretation of root infinitives and bare nouns in child language (=MIT Working Papers in Linguistics # 12)*. Cambridge, MA: MIT.

Lasser, I. (2002). The roots of root infinitives: Remarks on infinival main clauses in adult and child language. *Linguistics*, *40*, 767–796.

Lewis, J. D., & Elman, J. L. (2001a). A connectionist investigation of linguistic arguments from the poverty of the stimulus: learning the unlearnable. *Proceedings of the Twenty-Third Annual Conference of the Cognitive Science Society, Mahwah, NJ: Lawrence Erlbaum Associates*, 552-557.

Lewis, J. D., & Elman, J. L. (2001b). Learnability and the statistical structure of language: Poverty of stimulus arguments revisited. *Proceedings of the 26th Annual Boston University Conference on Language Development*, *1*, 359–370.

Matthews, D. (Ed.) (2014). *Pragmatic development in first language acquisition*. Amsterdam: Benjamins.

McCauley, S. M., & Christiansen, M. H. (2019). Language learning as language use: A cross-linguistic model of child language development. *Psychol Rev*, *126*(1), 1–51. doi:10.1037/rev0000126

Mintz, T. H. (2003). Frequent frames as a cue for grammatical categories in child directed speech. *Cognition*, *90*(1), 91–117. doi:10.1016/s0010-0277(03)00140-9

Moran, S., Blasi, D. E., Schikowski, R., Kuntay, A. C., Pfeiler, B., Allen, S., & Stoll, S. (2018). A universal cue for grammatical categories in the input to children: Frequent frames. *Cognition*, *175*, 131–140. doi:10.1016/j.cognition.2018.02.005

Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.

Pinker, S. (1994). *The language instinct: How the mind creates language*. New York, NY and London, UK: William Morrow & Co (New York), Penguin (London).

Pinker, S., & Prince, A. (1991). Regular and irregular morphology and the psychological status of rules of grammar. *BLS*, *17*, 230–251.

Poeppel, D., & Wexler, K. (1993). The Full Competence Hypothesis of clause structure in early German. *Language*, *69*, 1–33.

Pullum, G. K., & Scholz, B. C. (2002). Empirical assessment of stimulus poverty arguments. *The Linguistic Review*, *19*, 9–50.

Ravid, D., Dressler, W. U., Nir-Sagiv, B., Korecky-Kröll, K., Souman, A., Rehfeldt, K., . . . Gillis, S. (2008). Core morphology in child directed speech: Crosslinguistic corpus analyses of noun plurals. In H. Behrens (Ed.), *Corpora in language acquisition research: Finding structure in data* (pp. 25–60). Amsterdam: Benjamins.

Ravid, D., Keuleers, E., & Dressler, W. U. (2020). Emergence and early development of lexicon and morphology. In V. Pirrelli, I. Plag, & W. U. Dressler (Eds.), *Word Knowledge and Word Usage* (pp. 593–633). Berlin, Boston: De Gruyter Mouton.

Reali, F., & Christiansen, M. H. (2005). Uncovering the richness of the stimulus: Structure dependence and indirect statistical evidence. *Cognitive Science*, *29*, 1007–1028. doi:http://dx.doi.org/10.1207/s15516709cog0000_28

Rumelhart, D. E., & McClelland, J. L. (1987). Learning the past tense of English verbs: Implicit rules or parallel processing? In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 195–248). Hillsdale, NJ: Erlbaum.

Schimke, S., & Dimroth, C. (2018). The influence of finiteness and lightness on verb placement in L2 German: Comparing child and adult learners. *Second Language Research*, *34*(2), 229–256. doi:https://doi.org/10.1177/0267658317723071

Schmid, H.-J. (2015). A blueprint of the Entrenchment-and- Conventionalization Model. *Yearbook of the German Cognitive Linguistics Association* (*CGLA*), *3*(1), 3–25. doi:10.1515/gcla-2015-0002

Schmid, H.-J. (2020). *The Dynamics of the Linguistic System: Usage, Conventionalization, and Entrenchment*. Oxford: Oxford University Press.

Slobin, D. I. (1985). Crosslinguistic evidence for the language-making capacity. In D. I. Slobin (Ed.), *The crosslinguistic study of language acquisition. Vol. 2: Theoretical issues* (pp. 1157–1249). Hillsdale, NJ: Erlbaum.

Slobin, D. I. (1997). The origins of grammaticizable notions: beyond the individual mind. In D. I. Slobin (Ed.), *The crosslinguistic study of language acquisition. Vol. 5: Expanding the contexts* (pp. 265–323). Mahwah, NJ: Erlbaum.

Szagun, G. (2011). Regular/irregular is not the whole story: The role of frequency and generalization in the acquisition of German past participle inflection. *Journal of Child Language*, *38*(4), 731–762. doi:10.1017/S0305000910000255

Szagun, G., Stumper, B., Sondag, N., & Franik, M. (2007). The acquisition of gender marking by young German-speaking children: Evidence for learning guided by phonological regularities. *Journal of Child Language*, *34*(3), 445–471. doi:dx.doi.org/10.1017/S0305000906007951

Theakston, A. L., & Lieven, E. (2017). Multiunit Sequences in First Language Acquisition. *Topics in Cognitive Science*, *9*(3), 588–603. doi:10.1111/tops.12268

Theakston, A. L., Lieven, E. V. M., Pine, J. M., & Rowland, C. F. (2002). Going, going, gone: The acquisition of the verb "go". *Journal of Child Language*, *29*, 783–811. doi:https://doi.org/10.1017/S030500090200538X

Tomasello, M. (1992). *First verbs: A. case study of early grammatical development*. Cambridge: Cambridge University Press.

Tomasello, M. (1998). The return of constructions. *Journal of Child Language*, *25*, 431–442. doi:https://doi.org/10.1017/S0305000998003493

Tomasello, M. (2000). First steps toward a usage-based theory of language acquisition. *Cognitive Linguistics*, *11*, 61–82.

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

Tomasello, M., & Brooks, P. (1998). Young children's earliest transitive and intransitive constructions. *Cognitive Linguistics*, *9*(4), 379–395.

Twomey, K. E., & Westermann, G. (2019). Building the foundations of language: Mechanisms of curiosity-driven learning. In J. Horst & J. von Koss Torkildsen (Eds.), *International handbook of language acquisition* (pp. 102–114). London: Routledge.

Welke, K. (2019). *Konstruktionsgrammatik des Deutschen. Ein sprachgebrauchsbezogener Ansatz*. Berlin, Boston: De Gruyter.

Westergaard, M. (2009). *The acquisition of word order: Micro-cues, information structure, and economy*. Amsterdam, Philadelphia: John Benjamins.

Wijnen, F., Kempen, M., & Gillis, S. (2001). Root infinitives in early Dutch child language: An effect of input? *Journal of Child Language*, *28*(3), 629–660. doi:10.1017/S0305000901004809

Wöllstein-Leisten, A., Hilmann, A., Stephan, P., & Vikner, S. (1997). *Deutsche Satzstruktur: Grundlagen der syntaktischen Analyse*. Tübingen: Stauffenburg.

Xanthos, A., Laaha, S., Gillis, S., Stephany, U., Aksu-Koç, A., Christofidou, A., . . . Dressler, W. U. (2011). On the role of morphological richness in the early development of noun and verb inflection. *First Language*, *31*(4), 461–479. doi:10.1177/0142723711409976

Yang, C., Crain, S., Berwick, R. C., Chomsky, N., & Bolhuis, J. J. (2017). The growth of language: Universal Grammar, experience and principles of computation. *Neuroscience and Biobehavioral Review*. doi:10.1016/j.neubiorev.2016.12.023