On the interplay between syntax and statistical learning on errors in argument structure*

Myung-Kwan Park · Euhee Kim**
(Dongguk University · Shinheung College)

Park, Myung-Kwan & Kim, Euhee. (2013). On the interplay between syntax and statistical learning on errors in argument structure. *The Linguistic Association of Korea Journal*, *21*(3), 55-78. It has often been reported (cf. Bowerman, 1974) that L1 children or L2 learners tend to make errors in the use of verbs or predicates. They over-passivize, over-causativize, or over-intransitivize verbs or predicates when they have insufficient or incorrect knowledge of the language they try to acquire or learn. We examine these three types of errors, investigating what part of syntax is responsible for these errors, and how language development or learning proceeds on the basis of the language input available to L1 children or L2 learners to overcome these errors. We propose that all three cases of over-generalization are ascribed to the functional category, i.e. the little *v*, which is responsible for argument structure alternations. L1 children or L2 learners are presumably lacking or deficient in this functional category. However, they capitalize on statistical learning to learn the exact category of a verb or predicate on the basis of its distributional properties, identifying what verb or predicate can combine with a right kind of little *v*.

Key Words: over-passivization, over-causativization, over-intransitivization, little *v*. argument structure, syntactic distribution, statistical learning

1. Introduction

L1 children or L2 learners often make errors in using a language especially

* We are grateful to the three anonymous reviewers of the journal for the valuable comments and suggestions.
* The first author is Myung-Kwan Park, and the second author, Euhee Kim.

in the course of acquiring/learning it. We can say that they do so because of insufficient or incorrect knowledge about the linguistic features of the language they try to acquire or learn. This paper is concerned in particular with errors that L1 children or L2 learners make regarding the use of predicates/verbs and their argument structure. More specifically, this paper is concerned with over-passivization as in (1), over-causativization (or over-(di)transitivization) as in (2), and over-intransitivization as in (3) below:

(1) **OVER-PASSIVIZATION**
a. The most memorable experience of my life was happened 15 years ago. (Zobl, 1989)
b. Rush hour traffic can be vanished because working at home is a new version. (Yip, 1995)

(2) **OVER-CAUSATIVIZATION**
a. I want to stay this rubber band on. (= let it stay on; leave it on) (Christy, 3;7)
b. Don't tight this 'cause I tight this. (= tighten) (Eva, 2;4)
c. I wanta…wanta…wanta round it. (= make it go around; turn it) (Christy, 3;0)
d. No, mommy, don't eat her yet, she's smelly! (= make her eat; feed her) (Christy, 3;8) (Bowerman, 1974)

(3) **OVER-INTRANSITIVIZATION**
a. How come these two broke? By who? (Christy, 7;6) (Bowerman, 1991)
b. I wanna take it out so it can't put on my nose. (Jennifer, 2;10) (Lord, 1979)

This paper explores these cases of errors from the two perspectives. One is syntax, and the other is statistical learning. Taking into account the three types of errors above, we look into what part of syntax comes into play, and to what extent statistical learning of relevant properties from the language input available to L1 children or L2 learners is instrumental in overcoming the errors in question.

2. Description of the data

The following examples show that L1 children do not make a clear distinction between passives and intransitive inchoatives. Based on her own diary notes of their spontaneous data, Bowerman (1991) reported that her daughters (Christy (C) and Eva (E)) used passives for spontaneous events for which adults would typically use intransitive inchoatives.

(4) a. E 5;2 Grasshoppers just jump in the grass so they can be hided.
b. E 6; 1 Do you wan to be died? (=*Do you want to die? In a fantasy game in which people die unless they take a certain medicine.*)
c. E 7;2 I wanta walk on a volcano, but one that's already been fired. I don't want to walk on one that's going to fire (=erupt; *note change to a more appropriate verb form in the second sentence.*)
d. C 6;9 They're not bloomed. (*Looking at cut flowers with buds that never developed.*) (p. 19)

Bowerman's daughters did not apply passives only to spontaneous events, either. She reported that her daughters also used intransitive inchoatives when they focused on what happened to the theme and left the agent argument DP out of consideration, where adults would prefer passives to intransitive inchoatives.

(5) a. C 2;3 It blowed up (=inflated; *Right after F inflated a beachball for C.*)
b. C 3;9 And then the cookie swallowed and (then) went down down down. (*Telling a story about the adventures of a cookie.*) (p. 24-5)

The use of oblique agent phrases *by* or *from PP* together with intransitive inchoatives also implies children's incomplete knowledge about the difference between passives and intransitive inchoatives.

(6) a. C 7;6 How come these two broke? By who? (*Holding pieces of a construction toy*)
b. E x;x It scrunched up from Brandon and Barclay.[1] (p. 24-5)

c. C 2;11 Bert knocked down. (*After seeing Bert topple over on TV.*)
d. C 4;5 But the parts might lose. (*Concerned about taking a game to school with her.*)

Lord (1979) also reported similar examples of intransitive inchoative errors, as follows[2]:

(7) a. Maybe it's a building building up. (Damon 2;9)
b. I wanna take it out so it can't put on my nose. (Jennifer 2;10)
c. Come and see what Jenny got today. Pull. Pull! (Benjy 3;1)
d. They didn't throw in! (Damon 3;4)
e. You keep on talking to her! And that makes me bother! (Benjy 3;11)

In the context of (7c) *pull* is erroneously used as an intransitive inchoative: the child is pulling on his reluctant mother and asking that she come along. And note that the inchoative verb *bother* in (7e) is the complement of the verb *make* in the periphrastic causative construction.

Bowerman (1974) also found errors in the opposite direction, with strictly intransitive inchoative verbs produced as transitive causatives in causative errors. The following examples, taken from Bowerman (1974), illustrate the over-causativization of an intransitive verb or even an adjective.

(8) a. C 3;5 How come you had a little troubling going it? (*M [Mother] couldn't start a car.*)
b. C 7;5 But he disappeared the green one and he disappeared the blue one! (*Watching magician do tricks with scarves on TV.*)
c. C 12;3 Salt clings it together. (*As C mixes playdough.*)
d. E 4;3 Can I glow him? (*Wants to play with a monster toy that glows*

1) Children use *from* instead of *by* at the early stage of development of passives (Bowerman, 1991).
2) Throughout the paper, the **Damon** errors are from Clark's (1993) unpublished diary data; the **Jennifer** and **Benjy** errors, from Lord (1979); and the rest, from Bowerman (1974, 1982, 1991).

after being held under a light.)
e. C 5;0 OK. If you want it to die. E's gonna die it. She's gonna make it die. (*C's sister E is about to touch a moth.*)
f. E 2;4 Don't tight this 'cause I tight this. (= tighten)

Likewise, Marcotte (2005) reported that children overgeneralize the causative alternation: they make causative alternation errors. Diary studies of children (Lord, 1979; Clark, 1993) also documented some more of these uses. For instance, the verbs in (9) (including (9a) and (9b), again taken from Bowerman (1974)) are erroneously used as transitive causatives in causative errors:

(9) a. Don't giggle me. (Eva 3;0)
b. I'm singing him. (Christy 3;1)
c. First I want to resquirt him and then I'll hop him onto the side. (Damon 3;6)
d. You can drink me the milk. (Jennifer 3;8)
e. The tiger will come and eat David and then he will be died and I won't have a little brother any more. (Hilary 4;x)[3)]

Note that (9d) is an erroneous ditransitive use of a transitive verb; we subsume such examples under the term transitive causative for expository purposes. Note also the passive use of transitive *die* in (9e).

3. The problem with Pinker s (1989) analysis

In this section, we are going to review Pinker's (1989) initial attempt to account for argument structure alternations between intransitive inchoatives and transitive causatives, and the L1 acquisition of them. On Pinker's (1989) account, linking rules for argument structure alternations are, in essence, semantic operations rather than syntactic ones, so as a natural consequence they discern the semantic features of verbs. In other words, correct linking must follow

3) The **Damon** errors are from Clark's unpublished diary data; the **Jennifer** and **Benjy** errors from Lord (1979); and the rest from Bowerman (1974, 1982).

automatically from meaning. Since Pinker's theory postulates that children are born with linking rules, children must be able to link the arguments of a verb correctly as long as children have acquired the meaning of the verb correctly.

For example, the rule for the causative alternation, in Pinker's analysis, takes a predicate that denotes a 'change' (an event of acting or moving in some way) and alters it into a predicate that denotes 'by acting on, cause to change (in a certain specified way).' (Or vice versa, this time deriving an intransitive inchoative verb from a transitive causative: the rule is bidirectional and can work in either direction.) Linking rules dictate that the first argument of CAUSE, the agent, is realized as the subject and the second argument, the affected entity, is realized as the object.

But there is a critical shortcoming in Pinker's account: there are verbs that meet these restrictions and yet do not alternate. For example, why can't *disappear* undergo the causative alternation, given that it satisfies the requirement that the verb denote a change? To cope with this problem, Pinker (1989) proposed breaking down each rule for alternation into two levels: a broad-range rule and one or more narrow-range rules, which are semantically more narrowed-down/finer-grained versions of the broad-range rule. The broad-range rule imposes the necessary conditions for a verb to alternate, but does not stipulate whether or not it actually does alternate. The narrow-range rules, in contrast, define the sufficient conditions.

In more detail, Pinker defined the broad-range rule for the causative alternation as restrictively as the facts of adult English will permit: First, the caused event must be <+dynamic> (i.e., the verb must have the feature ACT or GO in its semantic representation); second, the causing event must entail an ACT whereby an agent acts on a patient; and third, this act must bring about the caused event directly. Pinker moved on to propose that each broad-range rule is associated with one or more narrow-range rules: from the candidate alternating verbs allowed by the broad-range rule, the narrow range rules imposes a more fine-grained filter, selecting semantically homogeneous subclasses of verbs that do indeed alternate.

In view of language acquisition, children are expected to define the broad-range rule through a top-down process of formulation over verbs that have been observed to display the alternation. At the same time, they are expected to define narrow-range rules through a conservative bottom-up process in which the right to undergo the alternation generalizes, but only within the range of each semantic class for which an instance of an alternating verb has been observed. Crucially, what consists in a relevant semantic class of alternation is circumscribed by innate mechanisms (See Pinker (1989, p. 273-280)). The important thesis of Pinker's model is, then, that the child's rules are correctly constrained from the beginning,[4] so there is no need to explain how retreat or recovering from errors in argument structure alternations can occur in the absence of negative evidence.

Second, a number of such errors as the ones we saw in section 2 are not really compatible with the view that the child's grammar is perfectly adult-like from the beginning except for one-shot innovations, which make a creative impromptu use of the broad-range rule, perhaps especially under communicative needs when the child doesn't know or can't remember a better verb. The frequency of many of the errors made by children, however, argues against Pinker's 'one-shot innovation' hypothesis (as Pinker (1989, p. 325) also recognizes). As Bowerman and Croft (2007, p. 289) note, for example, Christy (Bowerman's daughter) causativized *stay* (e.g., *stay* the door open) at least 43 times between the ages of 2;4 and 10;4, long after she knew and usually used the more appropriate verbs *keep* and *leave*. She also causativized *go* at least 28 times between the ages of 2;8 and 7;11, long after she knew verbs like *send* and *take*. Often the child causativized a verb erroneously even well after an appropriate counterpart for it already settled down in their speech (e.g., causative *come* vs. *bring*).

---

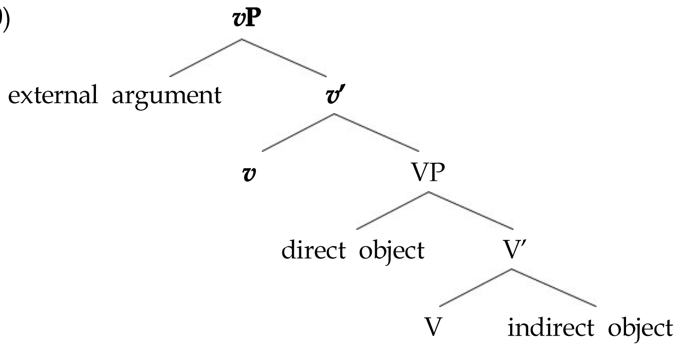
4) As reviewer C pointed out, Pinker (1989, 349) speculated that "an adult's narrow-class rules correspond to the verbs that happen to alternate in his . . . lexicon at a maturationally determined critical point, presumably around puberty". Now the question is how children come to formulate narrow-class rules up to puberty. They can do so relying on the semantics, or syntax, or semantics-syntax interaction of predicates. Refer to Ambridge, Pine, Rowland, and Young (2008) for the relevant discussions on this issue.

4. Toward an analysis

4.1. The development of the little *v*

Banking on the previous work by Hale and Keyser (1993), among others, Chomsky (1995) redefined the higher shell structure for VP proposed by Larson (1988) by suggesting that the upper verbal shell is not projected from an empty head, but from a phonetically null 'light' verb *v*, as represented schematically in (10):

(10)



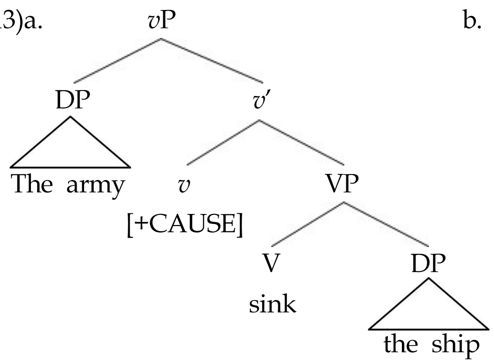
As Hornstein et al. (2005) pointed out, roughly speaking, a light verb is a verb whose meaning is heavily attributed to that of its complement. As can be seen in (11) below, the verb *take* in each of the sentences, for instance, is presumably rather different. This is because *take* in these sentences is a light verb and its meaning is contingent on the meaning of *shower* and *nap*. The light verb and its complement may thus be analyzed as forming a kind of complex predicate.

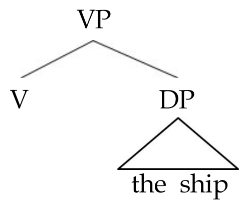
(11) a. John took a shower.
b. John took a nap.

Another oft-mentioned advantage of the double-shell structure for simple transitive constructions is that it accounts for the well-known relation between accusative Case and external theta-role, which renders itself into Burzio's (1986, p. 178) Generalization. According to this generalization, "[a]ll and only the verbs that can assign a Θ-role to the subject can assign accusative Case to an object. [subject = external argument]". Consider the causative/inchoative pair in (12), for example:

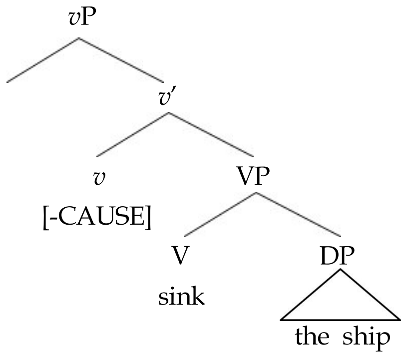
(12) a. Tom broke the cup. (Causative)
b. The cup broke. (Inchoative)

In (12a), the causative verb *break* assigns its external theta-role to *Tom* and accusative Case to *the cup*. In (12b), in contrast, the intransitive inchoative *break* does not assign an external theta-role, and neither does it Case-mark its object; *the cup* generated as a complement of the intransitive inchoative must then move to [Spec,TP] in order to be Case-marked. If simple transitive constructions also have two verbal shells and if the external argument is in the specifier of the outer shell, Burzio's Generalization may be regarded as a statement about the role of the light verb *v*: it is the element responsible for both external theta-role assignment and accusative Case-checking. Thus, the different properties of the causative/inchoative pair in (12) can receive an appropriate analysis if their verbal structures are accounted for along the lines of (13), with two shells for causatives as in (13a) and one for inchoatives as in (13b). Or alternatively, intransitive inchoatives may have the light verb, but this light verb carries the [-CAUSE] feature rather than [+CAUSE] feature, as in (13c):

(13)a.


b.


c.



Independent evidence for distinguishing causative/inchoative pairs in terms of verbal shells is taken from languages where the causative instance must have a verbal causative marker. In Kannada, for example, the causative version of (14a) contains the causative marker ***-is-***, as shown by the contrast between (14b) and (14c) taken from Lidz (2003):

(14) Kannada

a. neer kud-i-tu
   water boil-PST-3SN
   'The water boiled.'

b. *naan-u neer-annu kud-id-e
   I-NOM water-ACC boil-PST-1S
   'I boiled the water.'

c. naan-u neer-annu kud-**is**-id-e
   I-NOM water-ACC boil-**CAUS**-PST-1S
   'I boiled the water.'

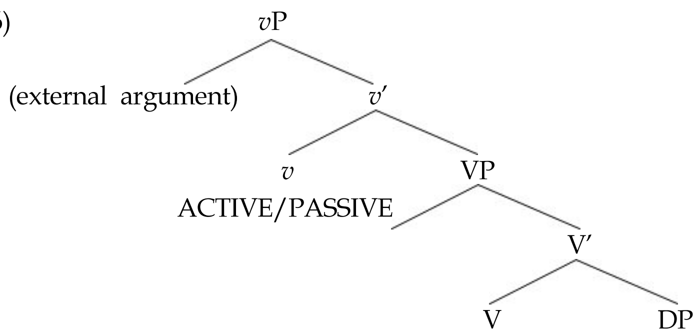
Given the analysis of (12) using the structure in (13), English and Kannada may be accounted for along the same line if ***-is-*** 'CAUSE' in (14c) is actually an overtly realized light verb, comparable to the phonetically empty little *v* in (13a).

The related point made by Hornstein et al. (2005) concerns active/passive pairs such as the one illustrated in (15).

(15) a. John built that house last year.
b. That house was built (by John) last year.

As is well known, passive constructions are taken to undergo a process suppressing accusative Case assignment and removing the status of the external theta-role by realizing it as an adjunct. If, as pointed out above, the proposed light verb of simple transitive constructions is the element that assigns both the external that-role and accusative Case, then it follows that a morpho-syntactic process over the light verb can change both Case- and theta-properties. Thus, the locus of passive and active alternations is ascribed to the little *v* and its properties, as represented in the following tree diagram:

(16)



The emerging generalization about argument structure alternations such as causatives and inchoatives, actives and passives, and transitives and intransitives is that they all come down to the little *v* and its properties. In other words, the little *v* can pick up features that determine the actual syntactic realization of both internal and external arguments in the verbal domain. We provisionally propose that the little *v* can be realized in several ways with different arrays of features, as follows:

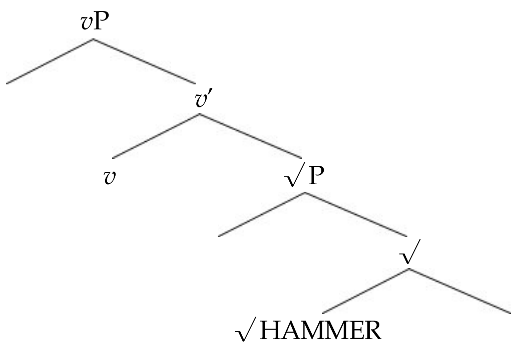
(17) a. causative *v* : [+Acc], [+EA]
b. simple intransitive *v* : [-Acc], [+EA]
c. inchoative *v* : [-Acc], [-EA], [-be$_{\mathbf{pass}}$]
d. passive *v* : [-Acc], [-EA], [+be$_{\mathbf{pass}}$]

Here [+EA] represents the requirement for assigning an external theta role, and [+$be_{\mathbf{pass}}$] represents the requirement for merging with a passive auxiliary verb *be*. The minus value represents the absence of such a requirement.

This role of the little *v* has a far-reaching consequence on the analysis of argument structure alternations. Since the little *v* is the locus of argument structure alternations, without making a right choice of it the final syntactic output will deviate from the correct form of the argument structure of a verb. As L1 children or L2 learners have the absence or limited use of functional categories in general,[5] it does not seem all that strange that they are prone to make errors in argument structure alternations, which we saw in the previous section.

Some remarks are in order on the category of a phrase selected by the little *v*. It has been known that the complement of the little *v* is the large VP. However, Marantz (1997, 2001) along with Arad (2003, 2005) proposed that this complement is projected from a category-neutral root, and the little *v* is a category determining head. For example, a verb like *hammer* consists of a root √HAMMER that could theoretically be either a noun or verb, and a functional head *v* that 'verbalizes' it.

(18)



5) Reviewer B commented that the kind of errors brought forth in this section that are made by L2 learners point to the fact that the L2 learner grammar in fact projects the little *v* in the clausal structure. In other words, L2 learners makes a productive use of it, even over-generalizing it to merge with inappropriate predicates. Obviously, in the text we didn't explore this possibility, but the possibility of children or learners starting with insufficient knowledge of but developing knowledge of the little *v* and relevant distributional properties in the course of learning or acquisition.

When we adopt this idea, it is not necessary to insert a verb as the head of the complement selected by the little *v*. Rather, it is possible to insert any category-neutral head that changes to a verbal little *v*P by combining with the category-determining little *v*. In fact, L1 children's over-causativization of adjectives or prepositions point to the fact that when they do not acquire the correct function of the little *v*, they can instead incorrectly use adjectives and prepositions for (causative) verbs, as can be seen by the following examples taken from Bowerman (1974):[6)]

Adjectives

(19) Don't **tight** this 'cause I **tight** this. (= tighten) (Eva, 2;4)
(20) **Full** it up! (= make it full; fill it up) (Christy, 2;3)
(21) I'm gonna **sharp** this pencil. (= sharpen) (Christy, 3;1)

Locative Particles

(22) I wanta…wanta…wanta **round** it. (= make it go around; turn it) (Christy, 3;0)
(23) **Up** your legs! (= make your legs go up; put your legs up) (Christy, 3;1)
(24) I'm gonna **round** it. (= make it round; roll it into a ball) (Eva, 3;8)

4.2. The lack/deficiency of the functional category *v* by native children or L2 learners

The inconsistent use of the functional morphemes and movement operations attributed to them in children's speech has raised questions about their role in early grammars. Some different proposals have been made, each taking either a Maturation or Continuity perspective. These two perspectives take different approaches to the nature of children's early syntactic representations and to the kinds of mechanisms inherent in developmental shift. From a Maturation standpoint, a child's grammar can have distinctive properties not present in an

6) Reviewer A claimed that there is a possibility that children use the null counterpart of the causative morpheme *-en* in (18)-(23). This means that a child's grammar inserts the [+CAUSE] feature in the little *v*, unlike the suggestion in the text that it lacks the functional category or relevant features constituting it.

adult system. Developmental changes in linguistic behavior are accounted for by qualitative shifts in the system underlying such behavior. In other words, changes in linguistic behavior stem from shifts in internal mechanisms. In contrast, from a Continuity standpoint, the child's grammar is molded of the same categories and principles as an adult grammar (Pinker, 1984). Developmental changes in linguistic behavior proceed in accordance with quantitative, incremental shifts in the existing system. External factors, or the interaction of external factors with the underlying system, can bring about developmental changes from a Continuity standpoint.

Proponents of one Maturation view have argued that no functional categories are projected in children's earliest grammars (e.g., Radford (1988, 1990, 1992)). They have suggested that the absence or limited use of functional morphemes and distributional traits attributed to them render evidence of deficits in underlying syntactic representation. According to Radford's (1988, 1990) Maturation account, children's early stage in syntactic acquisition maintains a lexical grammar, where utterances are constructed of lexical category projections, such as adjective phrase (AP), noun phrase (NP), verb phrase (VP), and prepositional phrase (PP), but lacking functional categories. Radford argued that utterances during this stage of child language are structurally analogous to adult small clauses rather than adult root clauses. As the name 'Maturation' suggests, researchers in favor of this position have proposed that the developmental shift introducing the projection of functional categories in children's grammars comes about through neurological maturation between the ages of 2;0 and 2;6.

From one Continuity standpoint, the child's initial representation of functional structure is assumed to consist of categories and principles found in the adult system. Advocates of less strong versions of continuity take omissions of functional morphemes and limited use of movement operations in children's speech possibly to point to deficits in syntactic representation. However, according to this perspective, changes in the underlying system throughout development do not come from a fundamental restructuring of competence through maturation, but instead follow from the gradual interaction of Universal Grammar (UG) and the ambient language input, which is instrumental in the emergence of a particular language grammar. One example of this view is the Functional Underspecification Hypothesis (Déprez, 1994). In order to account for the apparent optionality of DP-movement and V-to-C movement in early grammars, Déprez suggested that functional categories are fully available in children's syntax, but that certain parameter-related features of those categories could be underspecified at first, leading to representational/derivational differences between the adult's and the child's system. Similarly, Hoekstra, Hyams, and Becker (1997) suggested that the absence of morpho-syntactic markers of tense and definiteness in early grammars might be attributed to the initial underspecification of 'specificity' features in the nominal and verbal domains. Neither Déprez nor Hoekstra et al. offered an account of the mechanism that allows feature specification to occur eventually. Because they adopted a Continuity perspective, one could assume that sufficient exposure to the language input would eventually bring about the target feature specification.

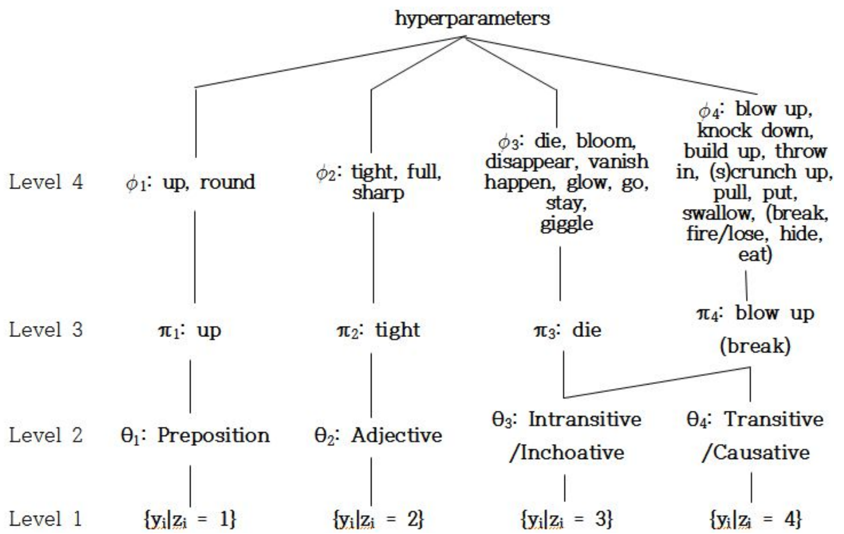
4.3. The development of the little *v* through statistical learning of argument structure from the relevant input

We adopt the Continuity perspective on the child's initial representation of functional structure, though it is immaterial in the following discussions which perspective we take. According to the Continuity perspective, changes in the underlying system throughout the child's L1 development are not brought about by a radical restructuring of competence via maturation, but instead results from the gradual interaction of Universal Grammar (UG) and the ambient language input, which molds a particular language grammar. As we saw in the section 4.1, UG dictates that a lexical item comes as category-neutral and the determination of its exact syntactic category is made by merging with the higher functional category such as the little *v*. Now the question is how the ambient language input comes into play in language development. We adopt the Hierarchical Dirichlet Process (HDP; Teh et al., 2006) model, which has been known to be effective in addressing this issue.

Like other topic models, the HDP model is in essence a model of category learning: this model groups together similar items like predicates in the input to learn structure. Embracing a usage-based approach to language (e.g., Goldberg, 2006), this model takes the acquisition of predicate argument structure to be a category learning task. In this analysis, structured knowledge turns itself into the hierarchical nature of the model.

The proposed model initially developed by Parisien and Stevenson (2010) adopts the HDP to address the acquisition or learning of predicate argument structure. The Figure (25) epitomizes the hierarchical levels of inference in this model. At level 1, the lowest level of abstraction, individual predicate usages y_i are depicted by sets of features[7].

(25) The proposed model: a Hierarchical Dirichlet Process applied to learning predicate argument structure constructions



At level 2, this model groups together similar usages to develop predicate argument structure constructions, where a construction is represented by a set of multinomial distributions[8], one for each feature. Since this clustering

7) We can assume such subcategorizational features as object, clausal complement, predicate complement, locative, adjunct, prepositional phrase, preposition, number of slots used, etc. We leave aside the issue of whether these features can derive from purely semantic/thematic features such as Agent, Theme, Proposition, etc (Grimshaw, 1979; Pesetsky, 1991).

mechanism is *nonparametric*, we need not presuppose the total number of constructions to cluster. Each of these constructions, denoted by its multinomial parameters $\theta = \theta_1, \theta_2, \dots$, probabilistically amounts to a pattern for a predicate such as a preposition, an adjective, an intransitive/inchoative verb, or a transitive/causative verb. While a construction here represents only syntactic information, with no semantic elements, this model can be generalized to a combined syntactic/semantic input representation.

At level 3, a multinomial distribution for each predicate, denoted by $\pi_p = \pi_1, \pi_2, \dots$, represents the range of constructions that tend to appear with a predicate. For example, in the Figure (25), *break* (π_4) would have a high probability both for the intransitive/inchoative and transitive/causative constructions (θ_3 and θ_4, respectively).

At level 4, we represent clusters of similar predicates. For each predicate cluster c_p, we use $\phi_{c_p} = \phi_1, \phi_2, \dots$, to represent the range of constructions that tend to occur with any of the predicates in that cluster. By serving as a prior on the predicate-level parameters π_p, ϕ_{c_p} directly influences each predicate in the cluster.

Let y_{ij} denote feature j of usage i. Levels 1 through 4 are given by the following:

(26)

$$\phi_{c_p} \sim \text{Dirichlet}(\alpha_0 \bullet \beta_0)$$ [9)]

8) In probability theory, the *multinomial distribution* is the one generalized from the binomial distribution. For *n* independent trials each of which results in a success for exactly one of k categories, with each category pertaining to a given fixed success probability, the multinomial distribution generates the probability of any particular combination of numbers of successes for the various categories.

9) In probability and statistics, the *Dirichlet distribution* (after Peter Gustav Lejeune Dirichlet), often represented as **Dirichlet(α)**, is an array of continuous multivariate probability distributions that are parametrized by a vector α of positive reals. It is the multivariate generalization of the beta distribution. Dirichlet distributions are very often employed as prior distributions in Bayesian statistics, and in fact the Dirichlet distribution is the conjugate prior of the categorical distribution and multinomial distribution.

Vlachos et al. (2009) employed a Dirichlet Process mixture model to group together verb

$$\begin{aligned} \pi_p &\sim \text{Dirichlet}(\alpha_1 \bullet \phi_{c_p}) \\ z_i &\sim \text{Multinomial}(\pi_p) \\ \theta_{jz_i} &\sim \text{Dirichlet}(1) \\ y_{ij} &\sim \text{Multinomial}(\theta_{jz_i}) \end{aligned}$$

Each predicate p pertains to a cluster of predicates c_p. Now, π_p depends on ϕ_{c_p}, which generates a distribution over constructions for all the predicates in the same cluster. The indicator variable z_i selects a cluster (i.e., a construction, one of the θ) for usage i. Given a predicate p, this is derived from a multinomial distribution which contains a small probability of introducing a new construction.

The predicate-specific distributions π_p and ϕ_{c_p} depend on hyperparameters α_0, α_1 and β_0 which denote expectations about constructions in general, across all predicates. They signify acquired knowledge about the likely total number of constructions, which constructions are more likely to be used overall, and so on:

(27)

$$\begin{aligned} \gamma_0 &\sim \text{Exponential}(1) \\ \alpha_{0,1} &\sim \text{Exponential}(1) \\ \beta_0 &\sim \text{Stick}(\gamma_0) \end{aligned}$$

β_0, drawn from a stick-breaking process (Stick)[10], denotes how many constructions will be employed and which constructions are more likely overall. α_1 influences the variability of π_p. α_0 influences the variability of ϕ_{c_p}. Large values of α_0 push ϕ_{c_p} closer to β_0, the global distribution over all the constructions in use, while smaller values result in more variation among predicates. γ_0 influences the total number of constructions overall; small values

types by their subcategorization preferences, but did not work out the issue of learning the argument structures themselves.

10) The so-called *stick breaking process* is the one that can be employed to make a constructive algorithm (the stick-breaking construction) for generating a Dirichlet process.

of γ_0 correspond to fewer constructions. By drawing $\alpha_{0,1}$ and γ_0 from an uniform exponential distribution[11], we give a weak preference for predicate-specific behaviour and for solutions with fewer constructions. These preferences are effectively designed into the model; they may be fine-tuned by general human category-learning behaviour. For further details of this model, see Teh et al. (2006).

To group predicates into alternation classes (see those predicates in (25) such as *break, fire, lose, hide eat*), we use a mechanism parallel to the way that we group individual predicate usages into constructions. Recall that c_p denotes an indicator variable, picking out a class for predicate p from the available classes in level 4. This is taken from a multinomial distribution σ which includes a small probability of introducing a new predicate class:

(28)

$$\gamma_1 \sim \text{Exponential}(1)$$
$$\sigma \sim \text{Stick}(\gamma_1)$$
$$c_p \sim \text{Multinomial}(\sigma)$$

As with earlier uses of the stick-breaking construction, γ_1 influences the expected total number of predicate classes. This method of clustering predicate types is tantamount to Wallach (2008).[12]

In essence, statistical learning of different types of predicates enables L1 children or L2 learners to decide which predicate can combine with a right kind of the little v. Classification of predicates is made by clustering them on the basis of their subcategorizational features in the language input available to L1 children or L2 learners. Note that the course of language development based on statistical learning makes a room for variations among language users in the

---

11) In probability theory and statistics, the *exponential distribution* (a.k.a. negative exponential distribution) is an array of continuous probability distributions. It depicts the time between events in a Poisson process, i.e. a process in which events occur continuously and independently at a constant average rate.

12) Wallace (2008) developed a hierarchical Bayesian model for clustering documents by topic. The model extended a well-known Bayesian topic model, latent Dirichlet allocation (Blei et al., (2003), to incorporate latent document groupings.

correct use of predicates. In other words, either when language users are not provided with sufficient language data or they make incorrect inferences in the course of clustering, they are apt to make errors in using predicates.

5. Summary and Conclusion

To account for the incorrect use of verbs or predicates by L1 children or L2 learners, we have proposed that the functional category in a clause, the little *v* plays an instrumental role in argument structure alternations. This functional category is responsible for the active/passive voice alternation, accounting for Burzio's generalization in a more systematic way. This means that if a certain verb does not have an external theta-role to assign (hence being unable to assign accusative Case), this verb cannot undergo passivization. There are also other peculiar restrictions on structure building with the little *v*. For example, it does not allow a certain apparently intransitive verb to merge with it, thereby blocking it from undergoing transitivization. On the other hand, the little *v* with the [-CAUSE] feature also does not allow a certain apparently transitive verb to merge with it, thereby preventing it from undergoing intransitivization. In a nutshell, the little *v* is the locus of argument structure alternations in human language.

We take the little *v* in clausal structure to be part of Universal Grammar. However, L1 children or L2 learners do not make consistent use of functional morphemes and functional categories. This means that they have insufficient or incorrect knowledge of the functional system in the language they try to acquire or learn. The absence or lack of knowledge about relevant features in the functional system of the little *v* leads to deviations from correct combinatorial composition of the little *v* with its complement predicate, thereby inducing over-passivization, over-causativization, and over-intransitivization.

We suggest that retreat or recovering from these types of errors involves learning of distributional contingencies for predicates. In other words, restrictions on the combination of the little *v* with a certain verb have to be learned from the language input available to L1 children or L2 learner. In particular, L1 children and L2 learners use statistics to identify particular abstract syntactic representations involving argument structure alternations (Miller and Chomsky, 1963; Pearl, 2007; Yang, 2006). According to this view, the learner may come equipped with a priori knowledge (UG), and the statistical learning interacts with that knowledge by determining the questions that the statistical distributions are relevant for answering, for the case at issue in this paper, the question of which predicate can combine with a right kind of little *v*. Hence, the outcome of the learning is a combination of the generalization emerging based on the observed input and innate knowledge.

References

Arad, M. (2003). Locality constraints on the interpretation of roots: The case of Hebrew denominal verbs. *Natural Language and Linguistic Theory*, 21(4), 737-778.

Arad, M. (2005). Roots and patterns: Hebrew Morpho-syntax. *Studies in natural language and linguistic theory*. Dordrecht: Springer.

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

Blei, D. M., Ng, A. Y., & Jordan, M. I. Latent Dirichlet allocation. *Journal of Machine Learning Research*, 3, 993-1022.

Bowerman, M. (1974). Learning the structure of causative verbs: A study in the relationship of cognitive, semantic, and syntactic development. In *Papers and Reports on Child Language Development* 17 (pp. 142-178). Department of Linguistics, Stanford University.

Bowerman, M. (1982). Evaluating competing linguistic models with language acquisition data: implications of developmental errors with causative verbs. *Quaderni di Semantica*, 3, 5-66.

Bowerman, M. (1991). When a patient is the subject: sorting out passives, anticausatives, and middles in the acquisition of English. Unpublished manuscript. Max Planck Institute for Psycholinguistics, Nijmegen, the Netherlands.

Bowerman, M., & Croft, W. (2007). The acquisition of the English causative alternation. In Bowerman, M., & Brown, P. (Eds.), *Crosslinguistic perspectives on argument structure: Implications for learnability* (pp. 279-307). Hillsdale, N.J.: Lawrence Erlbuam.

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

Clark, E. V. (1993). *The lexicon in acquisition*. Cambridge: Cambridge University Press.

Deprez, V. (1994). Underspecification, functional projections and parameter setting. In Lust, B., Suner, M., & Whitman, J. (Eds.), *Heads, projections and learnability* (pp. 249-271). Hillsdale, N.J.: Lawrence Erlbaum.

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

Grimshaw, J. (1979) Complement selection and the lexicon. *Linguistic Inquiry,* 10(2), 279-326.

Hale, K., & Keyser, S. J. (1993). On the argument structure and the lexical expression of syntactic relations. In Hale, K. and Keyser, S. J. (Eds.), *The view from building 20: Essays in linguistics in honor of Sylvain Bromberger* (pp. 53-109). Cambridge, MA: MIT Press.

Hoekstra, T., Hyams, N., & Becker, M. (1996). The underspecification of number and the licensing of root infinitives. *Proceedings of the 20th BUCLD* (pp. 293-306). Somerville, MA: Cascadilla Press.

Hornstein, N., Nunes, J., & Grohmann, K. 2005. *Understanding minimalism*. Cambridge: Cambridge University Press.

Larson, R. 1988. On the double object construction. *Linguistic Inquiry* 19, 335-392.

Lidz, J. 2003. Causation and reflexivity in Kannada. In Dayal, V., & Mahajan, A. (Eds.), *Clause Structure in South Asian Languages* (pp. 93-130). Dordrecht, Kluwer.

Lord, C. (1979). "Don't you fall me down": Children's generalizations regarding cause and transitivity. In *Papers and Reports on Child Language Development* 17 (pp. 81-89). Department of Linguistics, Stanford University.

Marantz, A. (2001). Words. *WCCFL XX*, University of Southern California, February 23-25.

Marantz, A. (1997). No escape from syntax: don't try morphological analysis in the privacy of your own lexicon. *University of Pennsylvania Working Papers* *in Linguistics*, 4(2), 201-225. Department of Linguistics, U. of Pennsylvania.

Marcotte, J.-P. (2006) Causative alternation errors as event-driven construction paradigm completions. In Clark, E. V. and Kelly, B. F. (Eds.), *Constructions in acquisition* (pp. 205-232). Stanford, CA: CSLI.

Miller, G., & Chomsky, N. (1963). Finitary models of language users. In Luce, R. D., Bush, R. R., & Galanter, E. (Eds.), *Handbook of mathematical psychology*, Vol. II. New York: Wiley.

Parisien, C., & Stevenson, S. (2010). Learning verb alternations in a usage-based Bayesian model. In Ohlsson, S., & Catrambone, R. (Eds.), *Proceedings of the 32nd Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

Pearl, L. (2007). *Necessary bias in natural language learning*. Unpublished doctoral dissertation, University of Maryland, College Park, MD.

Pesetsky, David. 1991. *Zero syntax vol. 2: Infinitives.* Unpublished manuscript available online.

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

Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure*. Cambridge: The MIT Press.

Radford, A. (1988). Small children's small clauses. *Transactions of the Philological Society*, 86, 1-43.

Radford, A. (1990). *Syntactic theory and the acquisition of English syntax: The nature of early child grammars of English*. Oxford: Basil Blackwell.

Radford, A. (1992). The acquisition of the morphosyntax of finite verbs in English. In Meisel, J. M. (Ed.), *The acquisition of verb placement: Functional categories and V2 phenomena in language acquisition* (pp. 23-62). Dordrecht: Kluwer Academic Publishers.

Teh, Y. W., Jordan, M. I., Beal, M. J., & Blei, D. M. (2006). Hierarchical dirichlet processes. *Journal of the American Statistical Association*, 101(476), 1566-1581.

Vlachos, A., Korhonen, A., & Ghahramani, Z. (2009). Unsupervised and constrained Dirichlet process mixture models for verb clustering. In *Proceedings of the EACL Workshop on Geometrical Models of Natural Language Semantics*.

Wallach, H. M. (2008). *Structured topic models for language*. Unpublished doctoral

dissertation, Univ. of Cambridge.
Yang, C. (2006). *The infinite gift: How children learn and unlearn the languages of the world*. New York: Scribner.
Yip, V. (1995). *Interlanguage and learnability: From Chinese to English*. Amsterdam: Benjamins.
Zobl, H. (1989). Canonical structures and ergativity. In Gass, S. M., & Schachter J. (Eds.), *Linguistic perspectives on second language acquisition* (pp. 203-221). New York: Cambridge University Press.

Myung-Kwan Park
Department of English
Dongguk University
26, 3-ga, Pil-dong, Chung-gu, Seoul
Phone: 82-2-2260-3153
E-mail: parkmk@dgu.edu

Euhee Kim
Major of Software Development, Division of Computer Information
Shinheung College
117, Howon-1-dong, Uyjeongbu, Gyeonggi Province
Phone: 82-31-870-3114
Email: euhkim@shc.ac.kr

Received on June 28, 2013
Revised version received on August 23, 2013
Accepted on September 5, 2013