A Bayesian Simulation of Clause-Level Constructional Knowledge in Child Language Development: Active Transitives and Suffixal Passives in Korean

Gyu-Ho Shin and Seongmin Mun

1. Introduction

A usage-based constructionist approach assumes that development of linguistic knowledge, represented as clusters of form-function pairings (i.e., constructions; Goldberg, 1995), occurs by way of interactions between language input and more basic forces from cognitive-psychological factors (e.g., Ambridge et al., 2015; Ellis, 2002; Lieven, 2010). The issue is how we better capture developmental trajectories of children's linguistic knowledge based on the exposure that they receive. One recent trend that draws attention to researchers in this respect is computational modelling, which provides a good estimation of how learning occurs (e.g., Ambridge & Blything, 2016; Lupyan & Christiansen, 2002; Matusevych et al., 2016). In particular, emerging research supports the effectiveness of Bayesian inference for this kind of task (e.g., Alishahi & Stevenson, 2008; Bannard et al., 2009; Barak et al., 2016; Nguyen & Pearl, 2019; Perfors, Tenenbaum, & Regier, 2011; Xu & Tenenbaum, 2007). One caveat is that, because previous research is skewed heavily towards English for its investigation, it is uncertain to what degree the implications of these simulation studies are generalisable across languages in support of the core assumption of the usage-based constructionist approach.

The present study explores how Korean-speaking children formulate their knowledge about representative argument structure constructions involving a transitive event (active transitives and suffixal passives) as a function of input properties and (non-)linguistic forces through the lens of computational modelling. We conduct a Bayesian simulation that employs information about the frequency of the two construction types found in CHILDES (MacWhinney, 2000). Korean, an under-studied language in this respect, is an agglutinative, SOV language with overt case-marking. These structural cues allow scrambling of pre-verbal arguments if that reordering preserves the original intention with no ambiguity. Korean also permits omission of almost all sentential elements: as long as participants in an event are clearly identified in the context, a case marker or a combination of an argument and a case marker can be omitted with the basic

* Gyu-Ho Shin, Palacký University Olomouc, gyuho.shin@upol.cz; Seongmin Mun, Université Paris Nanterre, simon.seongmin.mun@gmail.com.

 Proceedings of the 45th annual Boston University Conference on Language Development, ed. Danielle Dionne and Lee-Ann Vidal Covas, 679-692. Somerville, MA: Cascadilla Press.

propositional meaning intact. Although some studies reported Korean-speaking children's acquisition of these constructions through behavioural experiments (e.g., Jin et al., 2015; Kim et al., 2017; Shin, 2020), we are not aware of any study that investigates children's developmental trajectories involving the two construction types with a focus on Korean child corpora in this way. Our findings are thus expected to shed light on how children shape clause-level constructional knowledge in Korean, which is typologically different from the major languages currently under investigation in this respect.

As for the two constructions, a canonical active transitive (1a) typically occurs with the nominative-marked agent, followed by the accusative-marked theme. The thematic roles of each argument are indicated by designated case markers: a nominative case marker (NOM) *-i/ka* (*-i* after a consonant) and an accusative case marker (ACC) *-(l)ul* (*-ul* after a consonant). The two arguments can be scrambled, comprising the theme-agent ordering (1b). A canonical suffixal passive (2a) occurs with the NOM-marked theme, followed by the dative-marked agent indicated by a dative marker (DAT) *-eykey/hanthey*. The verb carries dedicated passive morphology as one of the four suffixes: *-i-, -hi-, -li-,* and *-ki-* (under allomorphic distribution). This pattern can be scrambled, yielding the agent-theme ordering, (2b).

(1a) Active transitive: canonical

Mina-ka	Ciwu-lul	an-ass-ta.
Mina-NOM	Ciwu-ACC	hug-PST-SE[1]

'Mina hugged Ciwu.'

(1b) Active transitive: scrambled

Ciwu-lul	Mina-ka	an-ass-ta.
Ciwu-ACC	Mina-NOM	hug-PST-SE

'Mina hugged Ciwu.'

(2a) Suffixal passive: canonical

Ciwu-ka	Mina-hanthey	an-ki-ess-ta.
Ciwu-NOM	Mina-DAT	hug-PSV-PST-SE

'Ciwu was hugged by Mina.'

(2b) Suffixal passive: scrambled

Mina-hanthey	Ciwu-ka	an-ki-ess-ta.
Mina-DAT	Ciwu-NOM	hug-PSV-PST-SE

'Ciwu was hugged by Mina.'

With these in mind, we specifically ask how Korean-speaking children's developmental trajectories of the two construction types, along with language-

[1] Abbreviation: ACC = accusative case marker; DAT = dative marker; NOM = nominative case marker; PSV = passive suffix; PST = past tense marker; SE = sentence ender

specific properties, can be understood as a function of input properties and statistical learning by way of computational modelling.

2. Bayesian simulation

Bayesian inference assumes that humans keep updating their beliefs about an event, represented as probabilities, through accumulated observations and make inferences by way of the updated beliefs. The degree of belief about an event (i.e., posterior probability) is calculated jointly by the accumulated degree of conviction in a hypothesis which occurs before encountering the event (i.e., prior probability) and a conditional probability where the event would be observed given that the hypothesis is true (i.e., likelihood) (Pearl & Russell, 2001; Perfors, Tenenbaum, Griffiths, & Xu, 2011). This idea is formalised as the Bayes' theorem (3a), where *A* and *B* are independent events, $P(A|B)$ refers to the posterior probability, $P(B|A)$ the likelihood, $P(A)$ the prior probability, and $P(B)$ the marginal probability.

(3a) $P(A|B) = (P(B|A) * P(A)) / (P(B))$

Oftentimes, P(B) is less important in actual application because *B* is fixed due to a stronger focus on the effects of *A* on our beliefs (Kruschke, 2015). This gives us a simpler formula (3b), where the posterior probability is proportional to the likelihood times the prior probability.

(3b) $P\ (A|B) \propto P\ (B|A)\ *\ P(A)$

Amongst previous Bayesian-inference-based studies, Alishahi and Stevenson (2008) provide an important precedent for the current work. They addressed a Bayesian way of emergence and growth of English verb-argument constructions, the results of which largely resembled developmental aspects that English-speaking children manifest. They created artificial input as pairs of a sentential frame and the corresponding semantic description involving the frame on the basis of naturalistic caregiver input in CHILDES. These form-meaning pairs were inputted to an unsupervised Bayesian learning model to measure how the model displayed probability distributions in the formation of constructional clusters as learning proceeded. Results showed that, as the quantity of input increased over time, the Bayesian model was able to not only assign higher probabilities to frequent verbs within specific constructions to which they were mapped but also generalise this schematic knowledge up to a newly attested lexicon. What they revealed in this computational modelling is consistent with the major assumptions of the usage-based constructionist approach, providing support for the interplay of frequency effects and general learning mechanisms without positing domain-specificity in language development.

They offer two conceptual points that are highly relevant to this study's aim. One is the direct mapping of a sentential frame and its semantic description. This reflects the idea that the inseparability of form and meaning/function, conceptualised as a construction, is a core property of language (Goldberg, 1995). We thus create input as a combination of a constructional frame (a morpho-syntactic layer) and its meaning/function (a semantic-functional layer). The other is about how constructions exist in humans' cognitive space. They assumed that constructional knowledge creates clusters that share similar features in their syntactic-semantic properties, intertwined with probabilities about how likely they accord with or deviate from each other (cf. Goldberg, 2019). Following this point, we demonstrate the growth of knowledge about the constructional patterns as clusters in the given simulation environment, by showing how posterior probabilities of these patterns change per learning.

2.1. Methods[2]

2.1.1. Input composition

All the constructional patterns for a transitive event were included, with scrambling and varying degrees of omission manifested (Table 1; adapted from Shin, 2020). Because there is no Korean corpus of caregiver input paired with semantic-functional information, we generated an artificial set of input based on the characteristics of Korean caregiver input in CHILDES pertaining to these patterns. In order to focus exclusively on the development of knowledge about clause-level constructions themselves, independently of concrete lexical items, we devised a set of schematised input comprising two layers in a pair: a morpho-syntactic layer specifying formal properties of the pattern and a semantic-functional layer indicating thematic roles of arguments and functions of markers. Each element in these layers had an index from the left to the right to maintain information about canonicity in the input.

To illustrate, the canonical active transitive (4) started with a nominal (N) followed by *-i/ka*, which was linked to the pair of the agent and the nominative. It proceeded with another nominal followed by *-(l)ul*, which was associated with the theme-accusative pair, and finally a verb (V) denoting an action. Whereas real morphemes indicated markers and passive morphology,[3] N and V represented abstract syntactic categories for noun and verb, respectively. Here, we do not assume that a child has these abstract categories in mind from the outset, but rather conceptualise them as a *heuristic*—strategic and provisional knowledge which is acquired probabilistically through exposure—that a learner employs in the course of acquisition: a word with a marker stands for an entity, and a word at the end of a sentence refers to an action.

---

[2] The Python code for this simulation is found at: https://github.com/seongmin-mun/Project/tree/master/Children%E2%80%99s%20development/Code/BayesianModule

[3] In the creation of input, we did not consider allomorphy involving case-marking and passive morphology, assuming that the occurrence of allomorphy is evenly distributed. We acknowledge the possibility that one allomorph occurs more frequently than the others or that the degree of form-function mappings of individual allomorphs may be disproportionate. This remains as one limitation of our simulation work.

Table 1. Constructional patterns (with or without scrambling and omission of sentential components) for a transitive event in CHILDES

Construction		Example	Freq (#) Caregiver	Freq (#) Child
Canonical active transitive	No omission	Mina-NOM Ciwu-ACC hug	1,757	37
	no ACC	Mina-NOM Ciwu-~~ACC~~ hug	268	14
	no NOM	Mina-~~NOM~~ Ciwu-ACC hug	19	0
Scrambled active transitive	No omission	Ciwu-ACC Mina-NOM hug	51	0
	no NOM	Ciwu-ACC Mina-~~NOM~~ hug	0	0
	no ACC	Ciwu-~~ACC~~ Mina-NOM hug	6	0
Active transitive	agent-theme, no CM	Mina-~~NOM~~ Ciwu-~~ACC~~ hug	3	0
	theme-agent, no CM	Ciwu-~~ACC~~ Mina-~~NOM~~ hug	0	0
	undetermined, no CM	Mina-~~NOM~~ Ciwu-~~ACC~~ hug	0	0
	agent-NOM only[1)]	Mina-NOM hug	935	21
	theme-ACC only[1)]	Ciwu-ACC hug	1,938	25
	agent only, no CM[1)]	Mina-~~NOM~~ hug	53	1
	theme only, no CM[1)]	Ciwu-~~ACC~~ hug	1,155	30
	undetermined, no CM[1)]	Mina-~~NOM~~ hug	40	1
Canonical suffixal passive	No omission	Ciwu-NOM Mina-DAT hug-psv	2	0
	no DAT	Ciwu-NOM Mina-~~DAT~~ hug-psv	0	0
	no NOM	Ciwu-~~NOM~~ Mina-DAT hug-psv	0	0
Scrambled suffixal passive	No omission	Mina-DAT Ciwu-NOM hug-psv	1	0
	no NOM	Mina-DAT Ciwu-~~NOM~~ hug-psv	0	0
	no DAT	Mina-~~DAT~~ Ciwu-NOM hug-psv	0	0
Suffixal passive	theme-agent, no CM	Ciwu-~~NOM~~ Mina-~~DAT~~ hug-psv	0	0
	agent-theme, no CM	Mina-~~DAT~~ Ciwu-~~NOM~~ hug-psv	0	0
	undetermined, no CM	Ciwu-~~NOM~~ Mina-~~ACC~~ hug-psv	0	0
	theme-NOM only[1)]	Ciwu-NOM hug-psv	407	9
	agent-DAT only[1)]	Mina-DAT hug-psv	13	0
	theme only, no CM[1)]	Ciwu-~~NOM~~ hug-psv	20	0
	agent only, no CM[1)]	Mina-~~DAT~~ hug-psv	0	0
	undetermined, no CM[1)]	Ciwu-~~NOM~~ hug-psv	0	0
Ditransitive	recipient-DAT only[2)]	Ciwu-DAT give	234	5
SUM			6,902	143

Note. CM = case-marking. 1) does not involve canonicity as it is undeterminable with only one overt argument. Although 2) does not relate to a transitive event *per se* and does not count as a relevant pattern, we considered it here because the DAT is often used as an indicator of a recipient in the active and thus a potential competitor of the agent-DAT pairing in the passive.

(4) Example of input: canonical active transitive, no omission

Morpho-syntactic layer	N_1–i/ka_1	N_2–(l)ul_2	V_3
Semantic-functional layer	Agent_1–NOM_1	Theme_2–ACC_2	Action_3

Note again that we did not include concrete words attested in the caregiver input to control for the effect of lexical words on the simulation results and to better demonstrate the developmental aspects of the constructional patterns themselves in the cognitive space that we modelled.

2.1.2. Model training

The general learning algorithm for our Bayesian learner was similar to that of Alishahi and Stevenson (2008): adding a new input item to an existing group of constructions that had the most similar characteristics to the item. The degree of similarity was determined by the probability that the new item was close to the individual constructional patterns in the model. This process is formalised as (5): in order to find the best-matching construction, the model classified a new input item *nCx* as an existing construction type *eCx*, ranging over the indices of all the constructions in the model, with the maximum probability given *nCx*.

(5) $\text{Best Construction}(nCx) = \underset{eCx}{\operatorname{argmax}}\, P(eCx \mid nCx)$

The computation of $P(eCx \mid nCx)$ followed the Bayes' rule as in (3b) where the posterior probability $P(eCx \mid nCx)$ was proportional to the multiplication of the conditional probabilities associated with the existing construction types and the prior of the existing construction types.

The frequency information in Table 1 served as initial priors for the constructional patterns; as learning proceeded, information about the constructional patterns was updated by adding the number of the classified input to the classified patterns over the course of learning. To prevent the probability from converging upon zero, we adopted the Laplace smoothing technique (e.g., Agresti & Coull, 1998): the Laplace estimator added the value of 1 as the Laplace value to the original frequency value so that the probability of occurrence of each construction type did not become zero and thus incalculable.

For construction learning, we used transitional probability, namely, a series of conditional probabilities from the first item to the last item within a specific pattern. This reflects how children utilise linguistic input for learning—figuring out intended meanings and functions given a form provided (cf. Goldberg, 2019)—in an incremental fashion (e.g., Özge et al., 2019; Strotseva-Feinschmidt et al., 2019). To illustrate, as Figure 1 demonstrates, the transitional probability of the canonical active transitive with no omission of arguments and case-marking is obtained by the multiplication of the following probabilities: construction-initial N-*i/ka* pairing (a), construction-initial agent-NOM pairing given the construction-initial N-*i/ka* pairing (b), construction-medial N-*(l)ul* pairing given the construction-initial agent-NOM pairings (c), construction-medial theme-ACC pairing given the construction-medial N-*(l)ul* pairings (d), construction-final V given the construction-medial theme-ACC pairings (e), and construction-final action given the construction-final V (f).

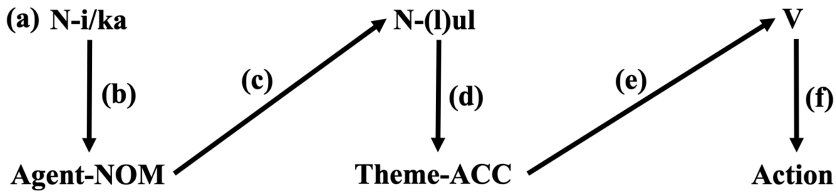

Figure 1. Schematic display of how to calculate transitional probability (canonical active transitive with no omission of arguments and case-marking)

2.1.3. Model performance and prediction

We set one learning phase for when all the input items (6,902 instances; see Table 1) were processed. Posterior probabilities of the constructional patterns were measured at every phase of learning (from one to ten) to estimate the degree of clustering for these patterns after the learning finished.

If the input characteristics outlined above influence our Bayesian learner, we should expect three outcomes. First, the degree of clustering for the constructional patterns should be asymmetric as learning proceeds. This asymmetry will be observed in the major increase in posterior probabilities of some dominant patterns frequently attested in the caregiver input (e.g., the canonical active transitive with no omission, the active transitive with only the theme-ACC pairing, the active transitive with only the theme argument without the ACC) relative to the other patterns. Next, the growth of clustering for the suffixal passive patterns should be suppressed throughout learning. Two possible factors will engage in this suppression effect: atypical case-marking (the NOM indicating the theme; the DAT indicating the agent) and unusual verbal morphology (attested less frequently than its active counterpart—the default form in our input). Third, as the input for our simulation did not include concrete lexical items, there should arise inconsistencies between the simulation results and the child production.

2.2. Results and discussion

Table 2 presents the posterior probabilities of the constructional patterns per learning. Whereas most of the constructional patterns converged upon almost zero probability, the canonical active transitive was the only pattern whose degree of clustering was constantly increasing as learning proceeded. The active transitive with only the theme-ACC pairing was the most frequent pattern for a transitive event attested in the input (1,938 instances), but the posterior probability of this pattern was neither the highest nor did it defeat that of the canonical active transitive with no omission. In contrast, the posterior probability of this the active transitive with only the no-ACC theme argument, the third most frequent pattern for a transitive event attested in the input (1,155 instances), increased until the fifth learning phase (with a small increase: 0.008) but it immediately decreased after then. It seems that the growth of the clustering for these patterns was somehow inhibited, possibly by the growth of the clustering for the fully-equipped canonical active transitive pattern during learning.

Table 2. By-construction posterior probability per learning

Type (example)	Posterior probability			
	1	3	5	10
Canonical active transitive, no omission	0.475	0.625	0.675	0.799
Scrambled active transitive, no omission	0.004	0.002	0.001	0.001
Canonical active transitive, no ACC	0.020	0.010	0.006	0.004
Canonical active transitive, no NOM	0.001	0.001	< 0.001	< 0.001
Active transitive, agent-NOM only	0.068	0.034	0.023	0.012
Active transitive, theme-ACC only	0.179	0.089	0.060	0.033
Active transitive, agent only, no case-marking	0.004	0.002	0.001	0.001
Active transitive, theme only, no case-marking	0.178	0.184	0.186	0.107
Active transitive, undetermined, no case-marking	0.002	0.001	< 0.001	< 0.001
Suffixal passive, theme-NOM only	0.029	0.015	0.010	0.005
Suffixal passive, agent-DAT only	0.001	0.001	< 0.001	< 0.001
Suffixal passive, theme only, no case-marking	0.002	0.001	< 0.001	< 0.001

Note. The other constructional patterns not listed in this table converged upon zero probability immediately after the 1st learning. The ditransitive pattern only with the recipient-DAT pairing does not fall into a transitive event, so we excluded the pattern in this table. For the sake of readers, it achieved the posterior probability of 0.035 and 0.036 after the 1st and the 10th of learning, respectively.

The degree of clustering for the remaining patterns decreased, the reason of which is ascribable to the same kind of suppression effects induced by their full-ledged constructional pattern (i.e., the canonical active transitive with no omission), which occupied a fairly large amount of input. Meanwhile, the reason that the posterior probability of the active transitive with only the agent-NOM pairing decreased over learning is somewhat unclear. We speculate that a similar kind of inhibitory force from various constructional patterns in the input affected how this pattern was learnt. This pattern occupied the fourth most frequent one appearing in the input. However, the agent-NOM pairing that occurred before a verb (935 instances for the active transitive, agent-NOM only; 6 instances for the scrambled active transitive, no ACC) was less frequent than the same pairing that occurred before the N-*(l)ul* pairing (1,938 instances for the canonical active transitive, no omission). This interplay may have suppressed the growth of this pattern.

The change of posterior probabilities in the passive patterns is attributable to the cue competition involving case-marking and verbal morphology. The suffixal passive with only the theme-NOM pairing has two features: the unusual case-marking (the NOM indicating the theme) and the atypical passive morphology. The growth of this pattern may have been suppressed greatly by its corresponding pattern—the active transitive with only the agent-NOM pairing, which has the typical case-marking (the NOM indicating the agent) and the typical verbal morphology (no active morphology). Similarly, the growth of the suffixal passive with only the agent-DAT pairing may have been constrained by the ditransitive with only the recipient-DAT pairing: case-marking (the DAT indicating the recipient is more frequent than the DAT indicating the agent) and verbal morphology (verb with no morphology is more frequent than verb with passive suffixes). The suffixal passive with only the no-NOM theme argument engages in passive morphology, which is atypical and possibly gives way to its similar composition—the active transitive patterns with only one case-less argument; 1,248 instances.

Whilst we found the global-level similarity between the model performance and the child production, there were some interesting inconsistencies between them (Table 3). Considering the overall number of the constructional patterns that they produced (143 instances; see Table 1), the children seemed to prefer the three patterns, all of which engage in the NOM, in production. In contrast, the learning model did not yield the corresponding rates of posterior probabilities for these patterns within the given simulation environment.

Table 3. Inconsistency between corpus analysis and the simulation

Type	Frequency of occurrence (corpus analysis)		Posterior probability (simulation; 10th learning)
	Caregiver input (#)	Child production (#)	
Active transitive, agent-NOM only	935	21	0.012
Canonical active transitive, no ACC	268	14	0.004
Suffixal passive, theme-NOM only	407	9	0.005

It seems that the model performance of the three patterns directly followed the characteristics of the caregiver input. The active transitive with only the agent-NOM pairing (935 instances) was outnumbered by the corresponding pattern with only the theme-ACC pairing (1,938 instances). This characteristic may have affected the posterior probability of the former, less frequent pattern through the raw frequency (935 instances vs. 1,938 instances). The canonical active transitive with no ACC (268 instances) was also less frequent than the fully-equipped counterpart (canonical active transitive: 1,757 instances), which may have influenced the posterior probability of the pattern through both the raw frequency (268 instances vs. 1,757 instances) and the transitional probability (the probability of N_2 given that of Agent_1−NOM_1 suppressed by the probability of Theme_2−ACC_2 given that of Agent_1−NOM_1). Likewise, the number of the suffixal passive with only the theme-NOM pairing (407 instances) was less than that of the active transitive with only the agent-NOM pairing (935 instances), and this may have guided the posterior probability of the passive pattern by way of both the raw frequency (407 instances vs. 935 instances) and the transitional probability (the probability of Theme_1−NOM_1 given N_1−i/ka_1 suppressed by the probability of Agent_1−NOM_1 given N_1−i/ka_1).

In contrast to the Bayesian learner, children in real life may employ the exclusive status of the NOM pertaining to a transitive event. The NOM was not only a very reliable cue to introduce the agent but also a very reliable outcome invited by the agent, and it occurred more frequently in the initial position than in the non-initial position (e.g., Shin, 2020). These characteristics may have led the children in CHILDES to primarily deploy the NOM to indicate the agent of a transitive event. Indeed, this interpretation is consistent with a line of behavioural research that shows early emergence of and heavy reliance on a heuristic that exclusively employs the NOM to indicate the agent role particularly at the initial word order slot in constructional patterns for a transitive event (e.g., Jin et al., 2015; Lee et al., 2013; Shin, 2020). If this way of thinking is valid, a learning model should seek to establish this NOM-related heuristic and reliably employ it in the learning process. Unfortunately, our model did not seem to demonstrate this aspect. Input properties for the model training could help explain this discrepancy. The caregiver input in CHILDES includes many partial and verb-less utterances, some of which involve various pairings of a noun and a marker. In the simulation, we utilised well-equipped instances (with at least one argument and a verb), thus ignoring these incomplete instances in the input. Our simulation thus cannot speak to this issue clearly. Subsequent research should incorporate information about the partial utterances into model training to see if model performance approximates the children's production tendencies more accurately when considering this additional information.

The case of the suffixal passive with only the theme-NOM pairing is still unclear. We speculate that two forces create inconsistencies in this pattern: the impact of partial utterances that we suggested earlier, and influences of lexical items. Of the nine instances of this pattern that the children produced, four included the verb *po-i-* 'see-PSV' and two included the verb *yel-li-* 'open-PSV' (and we could not find this way of skewness in the rest of the patterns that the children uttered). Therefore, the child production involving this pattern may have been limited to less-abstract and narrow-range schemata in its initial phase (e.g., Tomasello, 2003) Relatedly, as the child production in CHILDES was tied to specific tasks and contexts (e.g., question-and-answer, response to storybooks), the particular discourse in which they were situated may have affected their use of this pattern. We used neither content words nor discourse features in our simulation environment. Therefore, the model in this study did not capture these possibilities, which are left unaddressed in the current study and requires further investigation. Future research should thus verify the implications of this study from various angles, considering a more comprehensive set of input with lexical words attested in caregiver input, along with information about specific registers from which input is obtained.

3. Conclusion

Our simulation work provides somewhat different flavour in comparison to the previous research on this issue (e.g., Alishahi & Stevenson, 2008; Ambridge & Blything, 2016; Bannard et al., 2009; Barak et al., 2016; Matusevych et al., 2016). This is due to the two particular motivations of this study. One is that we modelled a child learner after the age of one or two, following the age range of the children in CHILDES. This is why we employed frequency information in the caregiver input as the initial priors of our learning model, instead of creating a *tabula rasa* model from scratch, with the assumption that our Bayesian learner already had varying degrees of prior probabilities involving the constructional patterns. The other motivation was that we intended to model the development of linguistic knowledge about clause-level constructions themselves. This led us to devising a set of schematised input with N and V as heuristics, instead of using concrete lexical words, for the model training. This is supported by the idea of the early emergence of abstract knowledge (and yet still requiring considerable amount of exposure for the maturation of that knowledge) advocated by an early abstraction account (e.g., Dąbrowska & Tomasello, 2008; Rowland et al., 2012; Saffran et al., 1996).

We acknowledge that, possibly due to these motivations and the particularities for implementing the motivations into the simulation, our computational model may not have performed exactly like humans in the target task, as shown in the child production. In particular, the fact that we did not compose input with concrete lexical items attested in the caregiver input appears to make the model impossible to capture this lexically-tied factor to the extent that human learners do when acquiring constructional knowledge, as in the case of the suffixal passive with only the theme-NOM pairing. In this respect, we admit that our answer to the question of how children's acquisition of the two construction types in expressing a transitive event can be understood as a function of input properties and statistical learning can only be partial at this point.

Nonetheless, we did find some nice compatibility of model performance with the child production. The distributional properties of the constructional patterns for a transitive event and the particular characteristics of form-function associations involving case-marking dedicated to these constructions found in the caregiver input yielded model performance that was largely consistent with the child production, despite no support from concrete lexical items. This, we believe, provides another piece of empirical evidence for the major tenet of the usage-based constructionist approach, also ensuring the status of argument structure constructions, independent of individual lexical items, as a psychological reality employed in human language behaviours (cf. Goldberg, 2019). Although this study is somewhat limited in precisely pinpointing the locus of the dissimilarity amongst the caregiver input, the child production, and the model performance (cf. Table 3), our approach to revealing learner's representations of abstract constructional knowledge through computational modelling paves the way for the empirical investigation of child language development regarding this issue in lesser-studied languages.

Speaking of the individual markers dedicated to the active transitive and the suffixal passive, the earlier literature has placed more emphasis on the NOM and the ACC in the active compared to the NOM and the DAT in the passive. A good deal of research explored how Korean-speaking children employ the NOM and the ACC in the active (e.g., Cho, 1982; Jin et al., 2015; Lee et al., 2013). Surprisingly, very few studies address the developmental aspects of the NOM and the DAT in the passive. Considering this trend in research, this study's findings from corpus analysis and computational modelling complement the niches about the linguistic environments with which Korean-speaking children are normally surrounded in light of form-function mapping of case-marking dedicated to argument structure constructions.

More broadly, our findings should be further verified and re-assessed with behavioural experiments, particularly on comprehension and/or processing of clause-level constructions. We compared the model performance with the child production; one possible caveat in this way of comparison is that the mode of outcome from the children (i.e., production) may serve as a confounding factor in the comparison, which requires additional caution when interpreting the results. Relative to the active employment of online measurement of children's sentence processing in major languages under investigation (e.g., Abbot-Smith et al., 2017; Huang et al., 2013; Özge et al., 2019; Strotseva-Feinschmidt et al., 2019), the processing-based research on child language in Korean is in its infancy. Furthermore, literature on Korean-speaking children's language development in consideration of language-specific properties at the level of clause-level constructions is thin (cf. Jin et al., 2015; Kim et al., 2017; Lee et al., 2013; Shin, 2020). Future work would thus benefit from exploring the extent to which the findings of computational simulations (with various learning algorithms) explain those from behavioural experiments on comprehension/processing of child language development in Korean. This is something that we plan to pursue next.

Despite this study's narrow scope of investigation (i.e., constructions involving a transitive event only) and its inherent limitations such as no inclusion of concrete lexical words, we believe the findings of this study to extend our current understanding of how linguistic knowledge about argument structure constructions in expressing a transitive event is organised in children's cognitive space as a function of input properties and domain-general learning capacities.

References

Abbot-Smith, Kirsten, Franklin Chang, Caroline Rowland, Heather Ferguson, & Julian Pine. (2017). Do two and three year old children use an incremental first-NP-as-agent bias to process active transitive and passive sentences?: A permutation analysis. *PloS one*, *12*(10), e0186129.

Agresti, Alan & Brent A. Coull. (1998). Approximate is better than "exact" for interval estimation of binomial proportions. *The American Statistician, 52*(2), 119–126.

Alishahi, Afra & Suzanne Stevenson. (2008). A computational model of early argument structure acquisition. *Cognitive Science*, *32*(5), 789–834.

Ambridge, Ben & Ryan P. Blything. (2016). A connectionist model of the retreat from verb argument structure overgeneralization. *Journal of Child Language, 43*(6), 1245–1276.

Ambridge, Ben, Evan Kidd, Caroline F. Rowland, & Anna L. Theakston. (2015). The ubiquity of frequency effects in first language acquisition. *Journal of Child Language, 42*(2), 239–273.

Bannard, Colin, Elena Lieven, & Michael Tomasello. (2009). Modeling children's early grammatical knowledge. *Proceedings of the National Academy of Sciences, 106*(41), 17284–17289.

Barak, Libby, Adele Goldberg, & Suzanne Stevenson. (2016). Comparing computational cognitive models of generalization in a language acquisition task. In Jian Su, Kevin Duh, and Xavier Carreras (Eds.), *Proceedings of the 2016 conference on Empirical Methods in Natural Language Processing* (pp. 96–106). Association for Computational Linguistics.

Cho, Sook Whan. (1982). *The acquisition of word order in Korean*. Unpublished Master's thesis. Department of Linguistics, University of Calgary.

Dąbrowska, Ewa & Michael Tomasello. (2008). Rapid learning of an abstract language-specific category: Polish children's acquisition of the instrumental construction. *Journal of Child Language*, *35*(3), 533–558.

Ellis, Nick C. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition, 24*, 143–188.

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

Goldberg, Adele E. (2019). *Explain me this: Creativity, competition, and the partial productivity of constructions*. Princeton, NJ: Princeton University Press.

Huang, Yi Ting, Xiaobei Zheng, Xiangzhi Meng, & Jesse Snedeker. (2013). Children's assignment of grammatical roles in the online processing of Mandarin passive sentences. *Journal of Memory and Language*, *69*(4), 589–606.

Jin, Kyong-Sun, Min Ju Kim, & Hyun Joo Song. (2015). The development of Korean preschooler' ability to understand transitive sentences using case-markers. *The Korean Journal of Cognitive and Biological Psychology, 28*(3), 75–90.

Kim, Shin-Young, Jee Eun Sung, & Dongsun Yim. (2017). Sentence comprehension ability and working memory capacity as a function of syntactic structure and canonicity in 5-and 6-year-old children. *Communication Sciences & Disorders*, *22*(4), 643–656.

Kruschke, John. (2015). Doing Bayesian data analysis: A Tutorial with R, JAGS, and Stan (2nd edition). London: Elsevier.

Lee, Woo Yeol, Min Ju Kim, & Hyun Joo Song. (2013) The development of Korean children's abilities to use structural cues for sentence comprehension. *The Korean Journal of Developmental Psychology, 26*(4), 125–139.

Lieven, Elina. (2010). Input and first language acquisition: Evaluating the role of frequency. *Lingua, 120*(11), 2546–2556.

Lupyan, Gary & Morten H. Christiansen. (2002, January). Case, word order, and language learnability: Insights from connectionist modeling. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 24. Retrieved from https://escholarship.org/uc/item/8nf95595

MacWhinney, Brian. (2000). The CHILDES project: Tools for analyzing talk (3rd edition). Mahwah, NJ: Lawrence Erlbaum.

Matusevych, Yevgen, Afra Alishahi, & Ad Backus. (2016). Modelling verb selection within argument structure constructions. *Language, Cognition and Neuroscience, 31*(10), 1215–1244.

Nguyen, Emma & Lisa Pearl. (2019). Using developmental modeling to specify learning and representation of the passive in English children. In Megan M. Brown & Brady Dailey (Eds.), *Proceedings of the 43rd Boston University Conference on Language Development* (pp. 469–482). Somerville, MA: Cascadilla Press.

Özge, Duygu, Aylin Küntay, & Jesse Snedeker. (2019). Why wait for the verb? Turkish speaking children use case markers for incremental language comprehension. *Cognition, 183*, 152–180.

Pearl, Judea & Stuart Russell. (2001). Bayesian networks. In Michael A. Arbib (Ed.), *The handbook of brain theory and neural networks* (pp. 157–159). Boston, MA: MIT Press.

Perfors, Amy, Joshua B. Tenenbaum, Thomas L. Griffiths, & Fei Xu. (2011). A tutorial introduction to Bayesian models of cognitive development. *Cognition*, *120*(3), 302–321.

Perfors, Amy, Joshua B. Tenenbaum, & Terry Regier. (2011). The learnability of abstract syntactic principles. *Cognition, 118*(3), 306–338.

Rowland, Caroline F., Franklin Chang, Ben Ambridge, Julian M. Pine, & Elena Lieven (2012). The development of abstract syntax: Evidence from structural priming and the lexical boost. *Cognition*, *125*(1), 49–63.

Saffran, Jenny R., Richard N. Aslin, & Elissa L. Newport. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926–1928.

Shin, Gyu-Ho. (2020). *Connecting input to comprehension: First language acquisition of active transitives and suffixal passives by Korean-speaking preschool children*. Unpublished doctoral dissertation. University of Hawai'i at Mānoa.

Sohn, Ho Min. (1999). *The Korean language*. Cambridge University Press.

Strotseva-Feinschmidt, Anna, Christine S. Schipke, Thomas C. Gunter, Jens Brauer, & Angela D. Friederici. (2019). Young children's sentence comprehension: Neural correlates of syntax-semantic competition. *Brain and Cognition, 134*, 110–121.

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

Xu, Fei & Joshua B. Tenenbaum. (2007). Word learning as Bayesian inference. *Psychological Review*, *114*(2), 245–272.

Proceedings of the 45th annual Boston University Conference on Language Development

edited by Danielle Dionne
and Lee-Ann Vidal Covas

Cascadilla Press   Somerville, MA   2021

Copyright information

ISSN 1080-692X
ISBN 978-1-57473-067-8 (2 volume set, paperback)

Ordering information

To order a copy of the proceedings or to place a standing order, contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, sales@cascadilla.com, www.cascadilla.com