Isolating second language learning factors in a computational study of bilingual construction acquisition

Yevgen Matusevych (Y.Matusevych@uvt.nl)
Department of Culture Studies, Tilburg University

Afra Alishahi (A.Alishahi@uvt.nl)
Department of Communication and Information Sciences, Tilburg University

Ad Backus (A.M.Backus@uvt.nl)
Department of Culture Studies, Tilburg University
PO Box 90153, 5000 LE Tilburg, the Netherlands

Abstract

The study of second language acquisition (SLA) is often hindered by substantial variability in the background of learners, their learning process and the input they receive. This diversity often makes it difficult to isolate specific learning factors and study their impact on L2 development. We present a computational study of SLA as an alternative methodological approach. By applying a usage-based computational model of construction learning on bilingual (German and English) input data, we analyze various learning variables in isolation. In particular, we investigate three factors: ratio between the amount of L1 and L2 input, age of L2 onset, and L2 frequency distribution. Our results are in line with experimental findings on the facilitatory effect of lower L1/L2 ratio and balanced L2 frequency distribution. We found no negative effect of later age of L2 onset on proficiency, which might be due to positive cross-linguistic transfer between German and English constructions.

Keywords: second language acquisition, computational cognitive modeling, learning factors, construction grammar

Introduction

Second language (L2) learners show substantial variation in their developmental trajectory and linguistic proficiency. A systematic study of Second Language Acquisition (SLA) must account for multiple factors regarding the characteristics of the *learner* (e.g., mother tongue or L1, learning abilities, motivation), the *learning process* (e.g., implicit vs. explicit, age and rate of L2 exposure, opportunities for language use) and the *linguistic input* (e.g., its qualitative and quantitative properties, linguistic distance between L1 and L2) (see, e.g., DeKeyser, 2013). The main obstacle for conducting such studies is finding large, homogeneous populations which allow for manipulating one factor at a time while controlling for the rest, a costly and at times impossible task.

In the light of these difficulties, it is surprising that until recently SLA has received little attention from the computational modeling community. Computational models can predict the impact of each learning factor on L2 development, and their prediction can suggest fruitful directions for human subject studies. In first language acquisition research, computational models have been widely used for studying the impact of the characteristics of input and the learning mechanisms on behavioral patterns and developmental trajectory, both at the level of words and constructions (see, e.g., Chater & Manning, 2006). However, there are only a few models

of SLA (see an overview by Li, 2013), most of which simulate learning the phonetic representations of words and their association with meaning (Li, 2009; Cuppini, Magosso, & Ursino, 2013, etc.) or with grammatical categories (Monner, Vatz, Morini, Hwang, & DeKeyser, 2013). To our knowledge, the only existing model that focuses on sentence structure and constructions in SLA is that of Rappoport and Sheinman (2005), but in fact, it learns only one language (see the section *Related computational studies* for a more detailed description of these models).

What is lacking is a computational framework which allows us to study the acquisition of linguistic constructions in L1 and L2 as an intertwined process, to systematically manipulate various types of learning variables mentioned earlier, and to study their impact on linguistic performance. In this paper, we take a first step in this direction by presenting a computational study of L1 and L2 construction acquisition, using an adaptation of an existing usage-based computational model (Alishahi & Stevenson, 2008). We report several experiments on how L2 proficiency is affected by a number of factors related to the learning process and the input.[1]

Learning factors

In the current study we manipulate three learning factors: the age of L2 onset, the ratio between the amount of L1 and L2 input, and frequency distribution in L2 input. The first two factors have been extensively studied in the SLA literature and shown to play a crucial role in L2 development, whereas experimental evidence on the impact of the third factor is inconclusive and at times contradictory.

L1/L2 Ratio. The amount and proportion of L2 input that a learner receives is often estimated by the amount of time they spend on learning and using L2. Muñoz (2011) reviews a number of studies showing that this factor correlates with learners' performance in various L2 proficiency tests.

Age of L2 Onset (*AO*). The impact of *AO* on L2 proficiency has been attributed to multiple sources, including biological constraints and the cognitive, socio-cultural and environmen-

---

[1]Preliminary results on a small data set were reported by Matusevych, Alishahi, and Backus (2013), see *Input data* for a comparison.

tal factors (see Han, 2004, for an overview). Within cognitive linguistics, the automatization of the L1 system (caused by L1 neural entrenchment) is believed to negatively affect L2 learning (e.g., MacWhinney, 2006).

Distribution of Input. It is subject to debate whether learning a new linguistic construction (both in L1 and L2) can be facilitated by a skewed frequency distribution of its instances in the input (for example, when one or two of the verbs which appear in a construction are much more frequent than others) (Boyd & Goldberg, 2009). While some studies demonstrate a positive effect of skewed input (e.g., Goldberg, Casenhiser, & Sethuraman, 2004), others do not (Year & Gordon, 2009), or even show quite the opposite (e.g., McDonough & Nekrasova-Becker, 2012).

Related computational studies

Perhaps the most fruitful computational research in bilingualism has been done using a family of self-organizing neural networks (see an overview by Li, 2009). Some experiments with these models study the pattern of lexical organization in L2 as a function of *AO*. These experiments show that the acquired L2 lexicon is independent and coexists with L1 lexicon in early L2 learning, but 'parasitizes' on L1 lexicon in late L2 learning (Zhao & Li, 2010). This difference is explained by decreased neural network plasticity and increased L1 entrenchment over time.

Monner et al. (2013) use a connectionist model to study potential impact of L1 entrenchment and memory development on L2 performance deficits. Their experiments also involve the manipulation of *AO*, and show a negative effect of L1 entrenchment on the acquisition of L2 grammatical gender. The model learns to associate each word form with the correct gender category over time.

Rappoport and Sheinman (2005) model the acquisition of L2 constructions, but their model suffers from inconsistency between L1 and L2 representations: L1 knowledge is static, in contrast to the emerging L2 system, which eliminates any possible effect of factors such as *AO* or L1/L2 ratio. Besides, they represent L1 only in terms of words and their relations, while the L2 system is purely syntactic – this discrepancy prevents cross-linguistic transfer of L1 forms into L2.

Computational model

In the current study we use an adapted version of an existing model of argument structure acquisition by Alishahi and Stevenson (2008). The model learns abstract constructions such as transitive and intransitive from instances of verb usage, and can use these constructions in a variety of language comprehension and production tasks. The model is informed by usage-based linguistics and is compatible with the Construction Grammar view that the building blocks of language are form–meaning pairings (Goldberg, 1995), but it proposes an alternative interpretation of linguistic constructions as a probabilistic association between lexical, syntactic and semantic features.

The example-based learning mechanism of Alishahi and Stevenson's (2008) model and its flexibility in simulating various aspects of language use makes it a suitable candidate for investigating how input-dependent factors (of the type discussed in the section *Learning factors*) affect SLA. In an SLA setup, the learner may receive input data which includes a mixture of usages from two different languages, although the underlying processing mechanism is the same. This is in line with views in cognitive linguistics, in particular the Unified Competition Model, which claims that the same cognitive mechanisms are used for L1 and L2 learning, and the difference between L1 and L2 proficiency levels can be attributed to factors such as L1 entrenchment (see MacWhinney, 2013).

Model architecture

We make the simplifying assumption that upon hearing an utterance in a perceptual context, the learner can recognize words in the utterance and also form a conceptualization of the described event. Furthermore, we assume that the learner knows the idiosyncratic meaning of words in the utterance and can identify their referent in the corresponding scene.

The model represents each individual verb usage as an *argument structure frame*, a collection of features that the learner can induce from the utterance and the observed event it refers to. These features include the predicate (corresponds to the main event) and its semantic properties, the number of arguments that the predicate takes, argument heads, their cases, their lexico-semantic and event-based (or thematic proto-role) properties, and finally the syntactic pattern (which in this case only reflects the order of arguments) and prepositions used in the utterance. In a bilingual setup, some features (i.e., semantic) share their range of values in L1 and L2, while other features (lexical) take language-specific values. A sample verb usage and its corresponding argument structure frame are shown in Table 1.

A construction represents a group of frames with similar features. By detecting and clustering similar frames, the model can abstract away from individual verb usages and form probabilistic associations between various features.

Table 1: An example frame extracted from a verb usage *I ate a tuna sandwich.*

predicate	*eat*
event properties	consume, take in, prep
arg. count	2
arg1	*I*
arg2	*sandwich*
arg1 lexical props	self, person, ..., entity
arg2 lexical props	snack food, dish, ..., entity
arg1 role props	living thing, entity, ..., organism
arg2 role props	solid, substance, ..., entity
arg1 case	N/A
arg2 case	N/A
syntactic pattern	ARG1 VERB ARG2
prepositions	N/A

Learning mechanism. The model receives one frame at a time and processes it using an incremental clustering algorithm which finds the most suitable construction for a new frame. This can be an existing construction or a new one:

$$\text{BestConstruction}(F) = \underset{k}{\arg\max}\, P(k|F) \qquad (1)$$

where k ranges over the indices of all constructions (index 0 represents the new construction). Using Bayes rule and dropping $P(F)$ which is constant for all k:

$$P(k|F) = \frac{P(k)P(F|k)}{P(F)} \propto P(k)P(F|k) \qquad (2)$$

The prior probability $P(k)$ indicates the degree of entrenchment of construction k:

$$P(k) = \frac{N_k}{N+1}, \; P(0) = \frac{1}{N+1} \qquad (3)$$

where N_k is the number of frames already contained in construction k, and N is the total number of frames observed so far. The posterior probability of a frame F is expressed in terms of the probabilities of its features, which are assumed to be independent:

$$P(F|k) = \prod_{i \in Features(F)} P(F_i|k) \qquad (4)$$

where F_i is the value of the ith feature in the frame F, and $P(F_i|k)$ is the probability of displaying the value F_i within construction k. This probability is estimated using a smoothed maximum likelihood formula. For single-valued features such as the head verb, likelihood is calculated by simply counting those members of construction k whose value for feature i exactly matches value F_i. For features with a set value such as the semantic properties of the verb and the arguments, the likelihood is calculated by comparing sets.

Using constructions. Linguistic productivity is simulated as predicting a missing value for a particular feature in a frame given the other available feature values. For example, verb comprehension can be seen as predicting the best value for the feature "event properties" given the lexical features (such as predicate and argument heads), the syntactic features (pattern and prepositions) and some semantic features (arguments' lexical and role properties). The probability that the missing feature i displays value F_i given other observed feature values in a partial frame F can be estimated as:

$$P(F_i|F) = \sum_k P(F_i|k)P(k|F) \propto \sum_k P(F_i|k)P(k)P(F|k) \qquad (5)$$

The probabilities $P(k)$, $P(F|k)$ and $P(F_i|k)$ are estimated as before (see equations 3 and 4). Ranging over the possible values of feature i, the value of an unobserved feature can be predicted by sampling possible values from the estimated $P(F_i|F)$ distribution.

Evaluating language proficiency. To evaluate the learner's language proficiency at a certain moment in time, we test the model on a language task where one of the features in a test frame is removed, and the model's prediction for the missing feature is compared to the original value. The *prediction accuracy* of feature i, or *PA_i*, is estimated based on the match between the original feature value and the value probabilities predicted by the model. For single-valued features such as frame predicate, this is estimated as the predicted probability of the original value. For features with a set value (e.g., event properties), this is estimated by measuring the Mean Average Precision[2] for the list of the predicted values ranked according to their probability, compared to the original set.

In this study, we test *PA_i* for three frame features – the predicate, its semantic properties, and arguments' role properties. Conceptually, predicting these features corresponds, respectively, to tasks of predicate selection, predicate comprehension, and interpreting the (thematic) role of each argument in the described event. In our preliminary experiments these three tasks proved to be most informative compared to others. For example, predicting syntactic pattern and prepositions appeared to be easy for the learner, due to the fact that probability mass functions for values of these features were strongly skewed towards a certain value. Finally, in the reported experiments we use a *language proficiency index* (*LPI*) defined as the average *PA_i* for the three ($n = 3$) features:

$$LPI = \frac{\sum_{i=1}^{n} PA_i}{n} \qquad (6)$$

Input data

In previous work (Matusevych et al., 2013), we compiled a small set of frames from usages of 6 German and 6 English verbs. The frames were manually extracted from German child-directed speech (L1) and English classroom material presented to German elementary school students (L2). Careful annotation of the conversations yielded a relatively noise-free data set, but the small size of our sample made the distribution of verbs and constructions unrealistic, and the outcome of our experiments unreliable.

For the experiments reported in this paper, we automatically extract frames for the same languages (German and English) from a number of available resources. Considering the frame features (see Table 1), we needed a syntactically parsed and POS-tagged corpus, also annotated with semantic role labels. For German, the TIGER corpus (Brants et al., 2004) appeared to be a good candidate, since it has been enriched with semantic role annotation in the SALSA project (Burchardt et al., 2006). For English, we used the WSJ part of Penn Treebank (Marcus et al., 1994), together with Proposition Bank (PropBank) containing argument structure information for WSJ (Palmer, Gildea, & Kingsbury, 2005). Although neither TIGER nor WSJ contained the kind of language that children or most L2 learners receive as input, we used these corpora as the only large sources for English and German that contained all annotations we needed.

From SALSA and PropBank, we extracted all verb-centered frames, each containing the main predicate (verb) and its arguments (labeled with their semantic roles and cases), as well as prepositions and word order. As for the

---

[2]Mean Average Precision is a standard measure in Information Retrieval, where a set of relevant items are expected to show up at the top of a ranked list of results.

latter, many German frames originated from verb-final embedded clauses, but in our setup frames represented simple main clauses, thus we had to recover German V2 word order by manually changing the predicate position.

We extracted the lexical properties of the noun arguments from WordNet (Miller, 1995). For verbs, we extracted features from VerbNet (Schuler, 2006) as well as WordNet. For adjectives and adverbs which are not hierarchically structured in WordNet, we used synonyms instead of hypernyms. To maximize semantic consistency, we used WordNet and Verb-Net for extracting the lexical properties of German words as well (by translating them into English). Finally, we manually compiled a set of lexical properties for frequent pronouns.

Arguments' role-based properties were also extracted from WordNet through an existing mapping between FrameNet semantic roles and WordNet (Bryl, Tonelli, Giuliano, & Serafini, 2012). This procedure resulted in two German and English data sets containing 3370 and 3803 frame instances, respectively. The data sets are comparable in terms of overall number of values for each feature, with the largest difference observed (obviously) for linguistic case, but also for argument role properties (353 for German vs. 195 for English) due to differences in granularity of the semantic annotated roles in SALSA and PropBank. In the experiments described below, we used both German and English as L1 or L2.

Experiments and results

In a simulation of an SLA scenario, the model receives and incrementally processes a sequence of N frames. Typically during the early stages of learning, all frames represent usages of L1. However, after a certain point in time (corresponding to our AO variable), the model receives a mixture of L1 and L2 frames, the proportion of which is determined by our Ratio variable (henceforth R). To depict the developmental trajectory of each language over time, we interrupt the training process at fixed intervals, and test the proficiency of the model on 20 test frames using the *LPI* measure (equation 6). Both training and test frames are sampled from the data sets described in the previous section. In the experiments reported here, we test the model at intervals of 20 frames, and the results are averaged over 10 simulations unless specified otherwise. Each simulation corresponds to a learner with a different history of linguistic input, and the 20 test frames differ for each learner as well.

Figure 1 shows L1 and L2 learning curves for two common scenarios: (a) early bilingual learners who receive equal amounts of L1 and L2 input from the start ($R = 1$, $AO = 0$); and (b) learners who are exposed to L2 input at a later stage and with a lower amount compared to L1 ($R = 5$, $AO = 200$). In both cases, the model receives a total of $N = 400$ frames. As can be seen, the L1 and L2 learning curves of the early bilinguals show the same developmental trajectory, whereas for the late L2 learners, the L2 curve straggles behind L1.

It must be noted that under equal conditions, learning English for our model is easier than learning German (as reflected in Figure 1(a)). This can be explained by the differ-
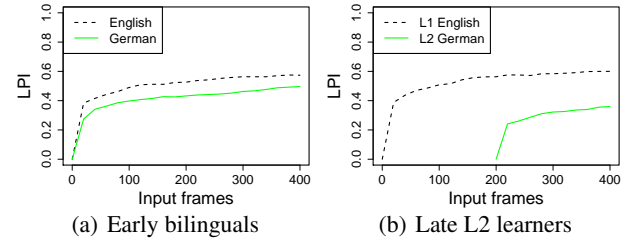

(a) Early bilinguals (b) Late L2 learners

Figure 1: *LPI* change over time in two specific settings

ence between English and German data sets – as mentioned before, the total number of values for arguments' role properties is larger in the German data set, which makes the task of predicting this feature more difficult in German than in English. However, a combination of larger R and AO can reverse this effect. In the late L2 learning scenario (Figure 1(b)) the lower level of L2 proficiency could be caused both by different R and AO values. The following experiments investigate the impact of each of the two variables in isolation.

L1/L2 Ratio

To investigate the impact of the proportion of L2 input data on proficiency, we ran a number of experiments with varying $R \in \{1, 2, 5, 10, 20\}$ while keeping the total number of frames and the age of L2 onset constant ($N = 400$, $AO = 0$). Figure 2 shows the developmental pattern (depicted by the change in L2 proficiency, or *LPI*) over time in varying R conditions. At the end of learning, *LPI* was negatively correlated with R both for L2 English (Kendall's tau $\tau = -.57$, $p < .001$) and for L2 German ($\tau = -.60$, $p < .001$). Figure 3 highlights this trend: the final *LPI* value significantly decreases as R increases.

This result is in line with experimental and observational findings that L2 proficiency is influenced by the amount of L2 exposure and contact with native speakers (see overview by Muñoz, 2011). However, there is no agreement yet on whether this effect is only limited to the initial period of learning. Although our results suggest a long-term effect, larger-scale simulations are needed to make a more reliable conclusion.
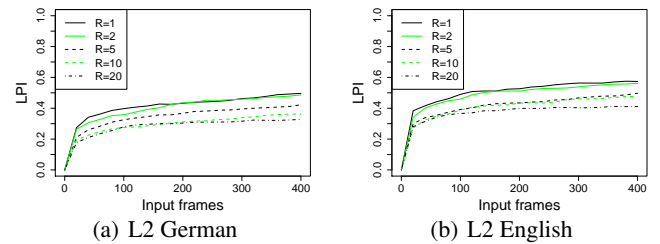

(a) L2 German (b) L2 English

Figure 2: L2 *LPI* change over time for different R values
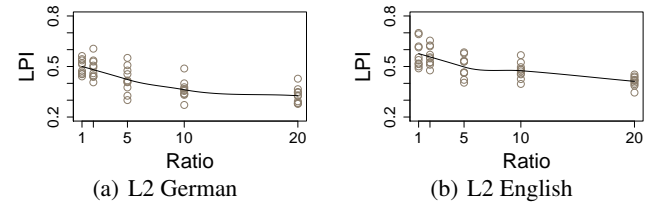

(a) L2 German (b) L2 English

Figure 3: Final level of L2 proficiency (*LPI*) in each simulation for different R values, with LOESS curves fitted

Age of L2 Onset

To study the impact of *AO* on L2 proficiency, we ran a number of experiments with varying $AO \in \{0, 100, 200, 300, 400\}$ (frames) while keeping R and N constant ($R = 1$, $N = 800$). For each *AO* value the model received the same amount of mixed L1/L2 input (hence the truncated curves). Figure 4 shows the change in L2 proficiency over time for different *AO* values. As we can see, there is no noticeable difference in the final *LPI* value across conditions. That is, we found no significant correlation between *AO* and the resulting *LPI*, either for English L2 ($\tau = .09, p > .05$) or German L2 ($\tau = .06, p > .05$).
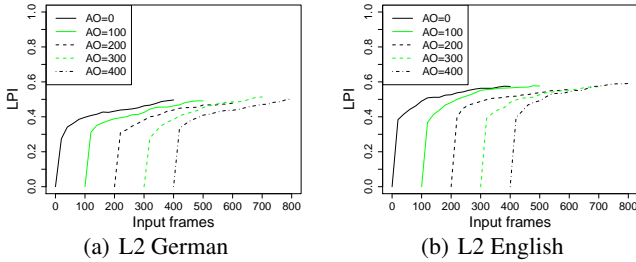
(a) L2 German   (b) L2 English

Figure 4: L2 *LPI* change over time for different *AO* values

These results are contrary to the L1 entrenchment account, which predicts that a more entrenched L1 uses up memory resources and yields lower L2 proficiency (Zhao & Li, 2010; Monner et al., 2013). However, considering the typological similarity between English and German, we believe that a positive transfer effect might be at play here, where L2 learning in higher *AO* conditions is facilitated by the existing knowledge of similar L1 constructions. Such effects must be observed for (shared) semantic features (e.g., event properties), but semantic features take set values, and it is difficult to trace the origin of each set element in the input. Nevertheless, the presence of transfer can be confirmed by looking at L2 predicate prediction. Predicting the L2 German predicate sometimes resulted in producing its L1 English counterpart (e.g., *occur* instead of *geschehen* [to occur], *increase* instead of *steigen* [to increase]), and vice versa (*existieren* [to exist] instead of *exist*). Thus, an absence of an *AO* effect might be due to the opposite directions of L1 entrenchment and positive transfer, and needs further investigation.

Frequency distribution in the input

To investigate whether the type of verb frequency distribution in L2 input affects L2 construction learning, we adopted a design from several experimental studies (e.g., Year & Gordon, 2009; Nakamura, 2012) where participants learn a new L2 construction which has no counterpart in their L1, and where the distribution of verbs in this construction is either skewed or balanced. Since it was difficult to find an English argument structure construction in our data set with no counterpart in German, in this experiment we only ran simulations on L1 English and L2 German. The construction of interest was a ditransitive with reversed order of arguments (THEME PRED. AGENT PATIENT, e.g., *das gab ich dem Herren* [I gave it to the gentlemen]), which was absent in the English data set. We manually prepared a small set of frames with 15 different
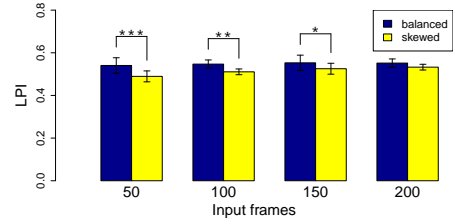
Figure 5: L2 *LPI* at 4 learning steps by the type of input

predicates appearing in this construction. In each simulation, 10 randomly selected predicates were used for training, and the other 5 for testing. In the balanced condition, train predicates were uniformly distributed, while in the skewed condition two randomly selected predicates were 20 times more frequent than the others. We ran the model with parameters $N = 400$, $AO = 200$, $R = 2$, and calculated LPI[3] (which in this case reflected the knowledge of a single construction) at 4 learning steps (50 frames each) after *AO*. The results (averaged over 30 simulations) are shown in Figure 5.

Applying Wilcoxon signed-rank test showed that after $n = 50$ mixed L1 and L2 frames, *LPI* was significantly higher for balanced (median $Mdn = 0.53$) than skewed input ($Mdn = 0.49, T = 46, p < .001, r = -.49$). The same effect was observed for $n = 100$ ($Mdn = 0.55$ vs. $0.52, T = 89, p < .01, r = -.38$) and $n = 150$ ($Mdn = 0.56$ vs. $0.52, T = 127, p < .05, r = -.28$), but there was no significant difference for $n = 200$ ($Mdn = 0.56$ vs. $0.53, T = 155, p > .05$).

The results support experimental findings that balanced input facilitates novel construction learning (Nakamura, 2012). McDonough and Nekrasova-Becker (2012), who found the same effect for an L2 construction that had a counterpart in the learners' L1, suggested that balanced input promotes broader category generalization. At the same time, other studies on learning a novel L1 construction found the opposite facilitatory effect of skewed input (see an overview by Boyd & Goldberg, 2009). Nakamura (2012) explains this mismatch by the fact that adult L2 learners, unlike L1 learners, engage in explicit learning. Our results support this argument, since our model performs an explicit categorization task by looking for regularities in instances, thus the learning is rather more explicit than implicit.

Conclusion

We present a computational study of learning second language constructions. Employing a usage-based computational model allows us to control the specifications of the learning process and the characteristics of input, and to simulate specific populations of L2 learners such as balanced bilingual children and late L2 learners. This approach enables us to isolate the impact of each learning factor on L2 development by manipulating one variable at a time – a methodological advantage that is absent from studies on human subjects.

Here we have investigated the impact of three learning variables: L1 to L2 ratio in the input, age of L2 onset, and L2 frequency distribution. Our experimental results showed that a

---

[3]In this experiment, learners never encountered test predicates in the input, thus instead of $PA_{predicate}$ we used $PA_{synt.pattern}$.

decreasing L1 to L2 ratio facilitates L2 development, as predicted by existing literature. Our analyses of L2 input distribution showed that balanced input facilitates construction learning, which is in line with some existing experimental findings. However, our simulations did not show a similar positive effect for early age of L2 onset. We suggest that this might be due to an interaction between two conflicting factors, L1 entrenchment and positive transfer from L1 to L2.

Our model represents both L1 and L2 as complex systems that comprise different features and can compete with each other. This framework allows for studying the concurrent acquisition of both languages, and modeling phenomena such as linguistic transfer. To our knowledge, this is the first modeling attempt of this kind. However, larger-scale and more diverse experimental investigation is needed for depicting a detailed picture of SLA and its parameters. In addition, future research must include other language pairs with different degrees of typological similarity in order to study the relative impact of cross-linguistic transfer.

Acknowledgements

References

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

Boyd, J. K., & Goldberg, A. E. (2009). Input effects within a constructionist framework. *The Modern Language Journal*, *93*(3), 418–429.

Brants, S., Dipper, S., Eisenberg, P., Hansen-Schirra, S., König, E., Lezius, W., . . . Uszkoreit, H. (2004). TIGER: Linguistic interpretation of a German corpus. *Research on Language and Computation*, *2*(4), 597–620.

Bryl, V., Tonelli, S., Giuliano, C., & Serafini, L. (2012). A novel Framenet-based resource for the semantic web. In *Proc. of SAC-2012* (pp. 360–365).

Burchardt, A., Erk, K., Frank, A., Kowalski, A., Pado, S., & Pinkal, M. (2006). The SALSA corpus: a German corpus resource for lexical semantics. In *Proc. of LREC-2006.*

Chater, N., & Manning, C. D. (2006). Probabilistic models of language processing and acquisition. *Trends in Cognitive Sciences*, *10*(7), 335–344.

Cuppini, C., Magosso, E., & Ursino, M. (2013). Learning the lexical aspects of a second language at different proficiencies: A neural computational study. *Bilingualism: Language and Cognition*, *16*, 266–287.

DeKeyser, R. M. (2013). Age effects in second language learning: Stepping stones toward better understanding. *Language Learning*, *63*, 52–67.

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

Goldberg, A. E., Casenhiser, D. M., & Sethuraman, N. (2004). Learning argument structure generalizations. *Cognitive Linguistics*, *15*(3), 289–316.

Han, Z. (2004). *Fossilization in adult second language acquisition*.

Li, P. (2009). Lexical organization and competition in first and second languages: computational and neural mechanisms. *Cognitive Science*, *33*(4), 629–664.

Li, P. (2013). Computational modeling of bilingualism: How can models tell us more about the bilingual mind? *Bilingualism: Language and Cognition*, *16*, 241–245.

MacWhinney, B. (2006). Emergent fossilization. In Z. Han & T. Odlin (Eds.), *Studies of Fossilization in Second Language Acquisition* (pp. 134–156).

MacWhinney, B. (2013). The logic of the unified model. In S. Gass & A. Mackey (Eds.), *The Routledge Handbook of Second Language Acquisition* (pp. 211–227).

Marcus, M., Kim, G., Marcinkiewicz, M. A., Macintyre, R., Bies, A., Ferguson, M., . . . Schasberger, B. (1994). The Penn Treebank: Annotating predicate argument structure. In *Proc. of the workshop on Human Language Technology* (pp. 114–119).

Matusevych, Y., Alishahi, A., & Backus, A. (2013). Computational simulations of second language construction learning. In *Proc. of CMCL-2013* (pp. 47–56).

McDonough, K., & Nekrasova-Becker, T. (2012). Comparing the effect of skewed and balanced input on English as a foreign language learners' comprehension of the double-object dative construction. *Applied Psycholinguistics*, *FirstView*, 1–24.

Miller, G. A. (1995). WordNet: A lexical database for English. *Communications of the ACM*, *38*(11), 39–41.

Monner, D., Vatz, K., Morini, G., Hwang, S.-O., & DeKeyser, R. M. (2013). A neural network model of the effects of entrenchment and memory development on grammatical gender learning. *Bilingualism: Language and Cognition*, *16*, 246–265.

Muñoz, C. (2011). Input and long-term effects of starting age in foreign language learning. *IRAL*, *49*(2), 113–133.

Nakamura, D. (2012). Input skewedness, consistency, and order of frequent verbs in frequency-driven second language construction learning. *IRAL*, *50*(1), 1–37.

Palmer, M., Gildea, D., & Kingsbury, P. (2005). The proposition bank: An annotated corpus of semantic roles. *Computational Linguistics*, *31*(1), 71–106.

Rappoport, A., & Sheinman, V. (2005). A second language acquisition model using example generalization and concept categories. In *Proc. of PsychoCompLA-2005* (pp. 45–52).

Schuler, K. K. (2006). *Verbnet: A broad-coverage, comprehensive verb lexicon*. Unpublished doctoral dissertation.

Year, J., & Gordon, P. (2009). Korean speakers' acquisition of the English ditransitive construction: The role of verb prototype, input distribution, and frequency. *The Modern Language Journal*, *93*(3), 399–417.

Zhao, X., & Li, P. (2010). Bilingual lexical interactions in an unsupervised neural network model. *International Journal of Bilingual Education and Bilingualism*, *13*(5), 505–524.